

Access DB# 76223

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval  
Reference Librarian  
Biotechnology & Chemical Library  
CM1 1E07 - 703-308-4498  
jan.delaval@uspto.gov

\*\*\*\*\*

### STAFF USE ONLY

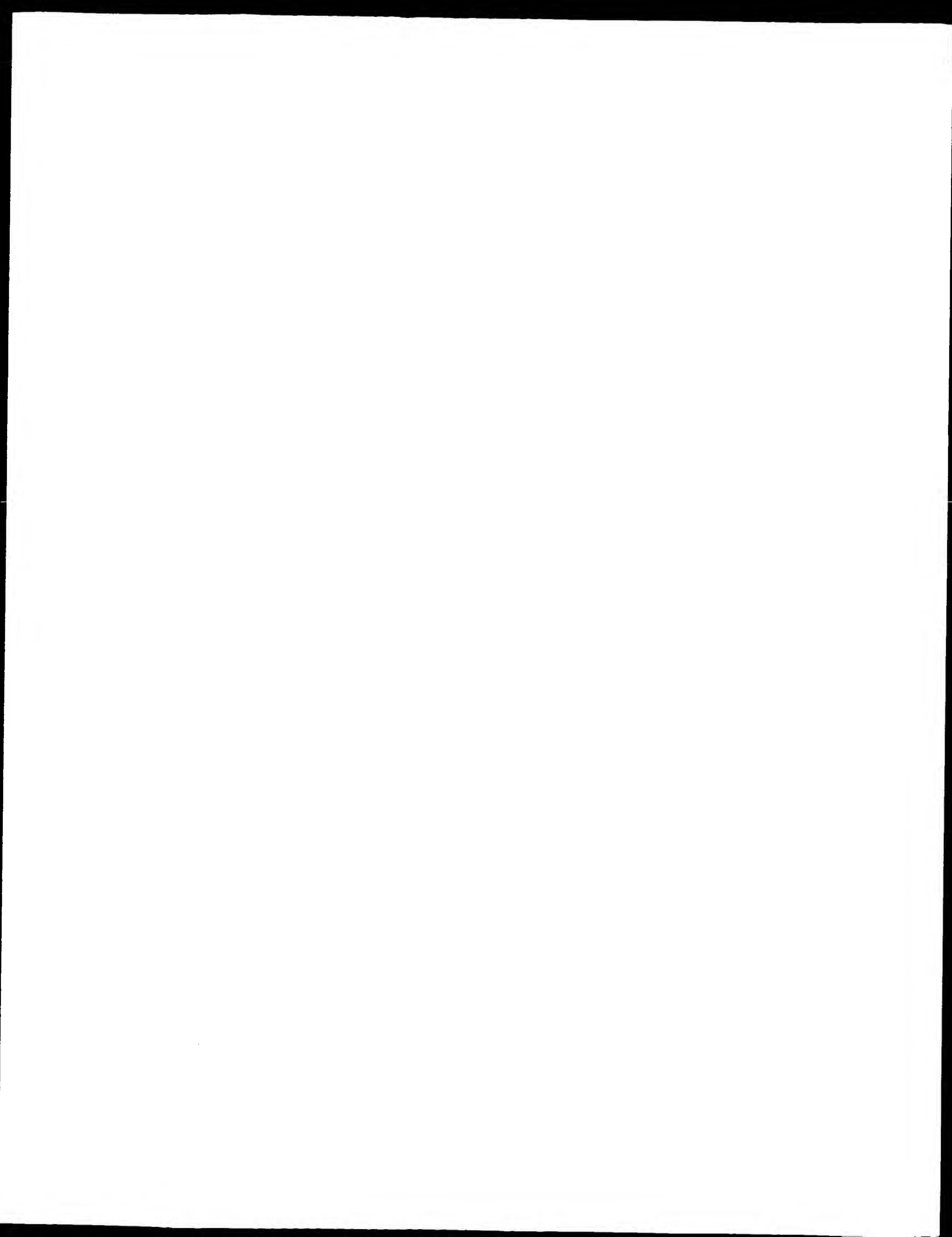
Searcher: Jan  
Searcher Phone #: 14498  
Searcher Location: 912102  
Date Searcher Picked Up: 9/22/02  
Date Completed: 9/22/02  
Searcher Prep & Review Time: 15  
Clerical Prep Time: 20  
Online Time: \_\_\_\_\_

### Type of Search

NA Sequence (#) \_\_\_\_\_  
AA Sequence (#) 7  
Structure (#) \_\_\_\_\_  
Bibliographic \_\_\_\_\_  
Litigation \_\_\_\_\_  
Fulltext \_\_\_\_\_  
Patent Family \_\_\_\_\_  
Other \_\_\_\_\_

### Vendors and cost where applicable

STN \_\_\_\_\_  
Dialog \_\_\_\_\_  
Questel/Orbit \_\_\_\_\_  
Dr.Link \_\_\_\_\_  
Lexis/Nexis \_\_\_\_\_  
Sequence Systems ☒ \_\_\_\_\_  
WWW/Internet \_\_\_\_\_  
Other (specify) \_\_\_\_\_





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:06:28 ; Search time 66.91 Seconds  
(without alignments)  
172.332 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_144  
Perfect score: 631  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....LDDFQKKQEMELXKQVE 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	631	100.0	267	1	LPBUA1	apolipoprotein A-I
2	606	96.0	267	1	A26529	apolipoprotein A-I
3	606	96.0	267	2	JS0079	apolipoprotein A-I
4	556	88.1	266	1	LPDGA1	apolipoprotein A-I
5	545	86.4	264	2	S31394	apolipoprotein A-I
6	545	86.4	265	2	A46018	apolipoprotein AI
7	533	84.5	265	2	JT0672	apolipoprotein A-I
8	521	82.6	265	2	A56858	apolipoprotein A-I
9	521	82.6	266	1	LPRB1Z	apolipoprotein A-I
10	517	81.9	265	1	LPRB1B	apolipoprotein A-I
11	516.5	81.9	241	2	A24998	apolipoprotein A-I
12	495	78.4	231	2	JQ0704	apolipoprotein A-I
13	481	76.2	262	2	JC1237	apolipoprotein A-I
14	481	76.2	264	2	S22420	apolipoprotein A-I
15	479	75.9	259	2	A24700	apolipoprotein A-I
16	329	52.1	264	1	LPCHA1	apolipoprotein A-I
17	328	52.0	264	2	JC5456	apolipoprotein A-I
18	326	51.7	246	2	A61448	apolipoprotein A-I
19	179	28.4	164	2	S21830	apolipoprotein A-I
20	142	22.5	429	2	S29565	apolipoprotein A-I
21	139	22.0	396	1	LPHUA4	apolipoprotein A-I
22	138	21.9	401	2	A47141	apolipoprotein A-I
23	125.5	19.9	391	1	LPRTA4	apolipoprotein A-I
24	114	18.1	391	2	B40892	apolipoprotein A-I
25	114	18.1	394	2	A25281	apolipoprotein A-I
26	114	18.1	395	2	A40892	apolipoprotein A-I
27	114	18.1	399	2	C40892	apolipoprotein A-I
28	106	16.8	34	2	S67972	apolipoprotein AI
29	106	16.8	36	2	A56866	apolipoprotein A-I

30	104	16.5	312	1	LPRTA	apolipoprotein E P
31	101.5	16.1	365	2	D71559	probable ABC ATPas
32	101	16.0	20	2	A05313	apolipoprotein A-I
33	98	15.5	298	2	S12635	apolipoprotein E P
34	98	15.5	311	2	JU0036	apolipoprotein E P
35	90	14.3	317	2	A28792	apolipoprotein E P
36	90	14.3	317	2	S03185	apolipoprotein E P
37	89	14.1	291	2	C60940	apolipoprotein E -
38	89	14.1	329	2	JC5566	apolipoprotein E P
39	89	14.1	1138	2	T24635	hypothetical prote
40	88	13.9	258	2	JH0472	apolipoprotein A-I
41	87	13.8	317	2	S33450	apolipoprotein E -
42	86.5	13.7	911	2	S51441	hypothetical prote
43	86.5	13.7	1046	2	T42720	cytoplasmic linker
44	86	13.6	336	2	T03999	AIg1 protein homol
45	86	13.6	583	2	C84788	probable myosin he

## ALIGNMENTS

RESULT 1  
LPBUA1  
apolipoprotein A-I precursor [validated] - human  
N;Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C;Species: Homo sapiens (man)  
C;Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C;Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010;  
6197  
R;Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A;Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A;Reference number: A90947; MUID:85026665  
A;Accession: A90947  
A;Molecule type: DNA  
A;Residues: 1-267 <SEI>  
A;Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
A;Accession: B90947  
A;Molecule type: mRNA  
A;Residues: 1-267 <SE2>  
A;Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R;Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988  
A;Title: Sequence and expression of Tangier apoA-I gene.  
A;Reference number: S02373; MUID:88196137  
A;Accession: S02373  
A;Molecule type: DNA  
A;Residues: 1-267 <MAK>  
A;Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
R;Shoulders, C.C.; Kornblith, A.R.; Munro, B.S.; Baralle, F.E.  
Nucleic Acids Res. 11, 2827-2837, 1983  
A;Title: Gene structure of human apolipoprotein AI.  
A;Reference number: A93465; MUID:83220822  
A;Accession: A93465  
A;Molecule type: DNA  
A;Residues: 1-267 <SHO>  
A;Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518;  
R;Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
A;Title: Isolation and characterization of the human apolipoprotein A-I gene.  
A;Reference number: A21147; MUID:84016011  
A;Accession: A21147  
A;Molecule type: DNA  
A;Residues: 1-267 <KAR>  
A;Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.  
Nucleic Acids Res. 12, 3917-3932, 1984  
A;Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundan  
A;Reference number: A93519; MUID:84221405  
A;Accession: A93519  
A;Molecule type: mRNA  
A;Residues: 1-267 <SHA>  
A;Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753

A;Accession: B93519  
 A;Molecule type: DNA  
 A;Residues: 1-24 <SH2>  
 R;Cheung, P.; Chan, L.  
 Nucleic Acids Res. 11, 3703-3715, 1983  
 A;Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
 A;Reference number: A93472; MUID:83220772  
 A;Accession: A93472  
 A;Molecule type: mRNA  
 A;Residues: 1-267 <CHE>  
 A;Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R;Law, S.W.; Brewer Jr., H.B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
 A;Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
 A;Reference number: A94010; MUID:84119464  
 A;Accession: A94010  
 A;Molecule type: mRNA  
 A;Residues: 1-267 <LAW>  
 A;Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R;Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
 A;Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
 A;Reference number: A21118; MUID:83195100  
 A;Accession: A21118  
 A;Molecule type: protein  
 A;Residues: 1-24 <ZAN>  
 R;Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
 Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
 A;Title: Human plasma proapo-A-I: isolation and amino-terminal sequence.  
 A;Reference number: A90112; MUID:83256553  
 A;Accession: A90112  
 A;Molecule type: protein  
 A;Residues: 19-27 <BRE>  
 R;Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
 Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
 A;Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high density lipoproteins  
 A;Reference number: A90209; MUID:78123731  
 A;Accession: A90209  
 A;Molecule type: protein  
 A;Residues: 25-57, 'Q', '59-169, 'QQ', '172-267 <BR2>  
 R;Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
 J. Clin. Invest. 82, 803-807, 1988  
 A;Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A-I)  
 A;Reference number: A30516; MUID:88331387  
 A;Accession: A30516  
 A;Molecule type: protein  
 A;Residues: 25-56 <YUI>  
 R;Nichols, W.C.; Dwulet, F.E.; Liepnieks, J.; Benson, M.D.  
 Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
 A;Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid.  
 A;Reference number: A31582; MUID:89050104  
 A;Accession: A31582  
 A;Molecule type: protein  
 A;Residues: 25-49, 'R', '51-85, 'D', '87-107 <NIC>  
 A;Note: variant sequence from patient with familial amyloidotic polyneuropathy type III  
 R;Manjunath, P.; Marcel, Y.L.; Uma, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A.  
 J. Biol. Chem. 264, 16853-16857, 1989  
 A;Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
 A;Reference number: A34409; MUID:89380318  
 A;Accession: A34409  
 A;Molecule type: protein  
 A;Residues: 25-48 <MAN>  
 R;Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
 A;Title: Structural requirements of human preproapolipoprotein AI for translocation and secretion  
 A;Reference number: S02737; MUID:89149957  
 A;Accession: S02737  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-6, 'AV', '9, 'LV', '12-29 <STO>  
 A;Note: part of this sequence, including the amino end of the mature protein, was confirmed  
 Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A;Title: Transient expression of wild type and mutant human apolipoprotein AI in COS cells  
 A;Reference number: S16197; MUID:92029676  
 A;Contents: annotation; extension of studies in reference S02737  
 R;Stoffel, W.; Kruger, E.; Deutzmann, R.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
 A;Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processed in vitro  
 A;Reference number: A19913; MUID:83236195  
 A;Accession: B19913  
 A;Molecule type: protein  
 A;Residues: 1-6, 'X', '8-13, 'XXX', '17-18, 'XX', '21, 'X', '23-25, 'X', '27-29 <ST2>  
 R;Ehnholm, C.; Bozas, S.E.; Tenkanen, H.; Kirsbaum, L.; Metsu, J.; Murphy, B.; Walke, B.  
 Biochim. Biophys. Acta 1086, 255-260, 1991  
 A;Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
 A;Reference number: A56815; MUID:92075698  
 A;Accession: A56815  
 A;Molecule type: protein  
 A;Residues: 25-31, 'P', '33 <EHN>  
 A;Experimental source: serum  
 A;Note: sequence extracted from NCBI backbone (NCBIP:69759)  
 A;Note: 32-Trp was also found  
 R;Kuntake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
 Biochemistry 33, 1988-1993, 1994  
 A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins  
 A;Reference number: A54223; MUID:94162201  
 A;Accession: A54223  
 A;Molecule type: protein  
 A;Residues: 25-39 <KUN>  
 R;Mogulievsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.; DNA 8, 429-436, 1989  
 A;Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: pAPOL  
 A;Reference number: I39476; MUID:89377481  
 A;Accession: I39476  
 A;Molecule type: mRNA  
 A;Residues: 19-267 <RES>  
 A;Cross-references: GB:M29068; NID:9178774; PIDN:AAA51747.1; PID:9178775  
 R;Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B.  
 J. Biol. Chem. 263, 18530-18536, 1988  
 A;Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by the promoter region  
 A;Reference number: I39475; MUID:89054040  
 A;Accession: I39475  
 A;Molecule type: DNA  
 A;Residues: 1-14 <RE2>  
 A;Cross-references: GB:J04066; NID:9178763; PIDN:AAA51746.1; PID:9553183  
 R;Breslow, J.L.  
 Annu. Rev. Biochem. 54, 699-727, 1985  
 A;Title: Human apolipoprotein molecular biology and genetic variation.  
 A;Reference number: A90042; MUID:85278004  
 A;Contents: annotation; review of sequences, variants and gene location  
 R;Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B.  
 J. Biol. Chem. 261, 3911-3914, 1986  
 A;Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation  
 A;Reference number: A92577; MUID:86140194  
 A;Contents: annotation; acylation with palmitate  
 A;Note: an undetermined serine or threonine is acylated by fatty acid; the acylating agent is not determined  
 R;Law, S.W.; Brewer, H.B.  
 J. Biol. Chem. 260, 12810-12814, 1985  
 A;Title: Tangier disease: The complete mRNA sequence encoding for preproapo-A-I.  
 A;Reference number: I55236; MUID:86008382  
 A;Accession: I55236  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-143, 'D', '145-267 <RE3>  
 A;Cross-references: GB:M11791; NID:9178776; PIDN:AAA35545.1; PID:9178777  
 C;Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestine  
 A;Title: HDL in plasma.  
 C;Genetics:  
 A;Gene: GDB:APOA1  
 A;Cross-references: GDB:119684; OMIM:107680  
 A;Map position: 11q23.3-11q23.3  
 A;Introns: 15/1; 67/2  
 C;Function:  
 A;Description: participates in the reverse transport of cholesterol from tissues to the liver  
 A;sterol acyltransferase (LCAT); noncovalently binds and stabilizes prostacyclin (PGI-2)

C:Superfamily: apolipoprotein A-I  
 C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipi  
 F:1-18/Domain: signal sequence #status experimental <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-267/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 100.0%; Score 631; DB 1; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-42;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
 |||||  
 Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELTYRQKVE 120  
 |||||  
 Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELTYRQKVE 144

## RESULT 2

A26529

apolipoprotein A-I precursor - crab-eating macaque

C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 30-Sep-1989 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999

C:Accession: A26529; A26627; S23135; A57766

R:Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marottili, K.R.

Gene 49, 103-110, 1986

A:Title: The primary structure of cynomolgus monkey apolipoprotein A-I deduced from the

A:Reference number: A26529; MUID:87191989

A:Accession: A26529

A:Molecule type: mRNA

A:Residues: 1-267 &lt;POL&gt;

A:Cross-references: GB:M15411; NID:g342074; PIDN:AAA36834.1; PID:g342075

R:Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolosi,

Biochemistry 26, 1457-1463, 1987

A:Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (cyn

A:Reference number: A26627; MUID:87185451

A:Accession: A26627

A:Molecule type: protein

A:Residues: 25-48 &lt;HER&gt;

R:Murray, R.W.; Marottili, K.R.

Biochim. Biophys. Acta 1131, 207-210, 1992

A:Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and corres

A:Reference number: S23135; MUID:92305062

A:Accession: S23135

A:Molecule type: DNA

A:Residues: 1-12, 'L', 14-267 &lt;MUR&gt;

A:Cross-references: GB:M63242; NID:g342070; PIDN:AAA36832.1; PID:g342071

R:Sorci-Thomas, M.; Kearns, M.W.

J. Biol. Chem. 266, 18045-18050, 1991

A:Title: Transcriptional regulation of the apolipoprotein A-I gene.

A:Reference number: A57766; MUID:92011532

A:Accession: A57766

A:Molecule type: DNA

A:Residues: 1-10 &lt;RES&gt;

A:Cross-references: GB:M69223; NID:g342066; PIDN:AAA36831.1; PID:g553820

C:Comment: The precursor is synthesized in the liver and small intestine. The propeptide

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HD

L) promoting cholesterol efflux from tissues and by acting as a cofactor for the lecithin

C:Genetics:

A:Introns: 15/1; 67/2

C:Superfamily: apolipoprotein A-I

C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe

F:1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-24/Domain: propeptide #status predicted &lt;PPT&gt;

F:25-267/Product: apolipoprotein A-I #status predicted &lt;MAT&gt;

Query Match 96.0%; Score 606; DB 1; Length 267;  
 Best Local Similarity 95.0%; Pred. NO. 4.7e-40;  
 Matches 114; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
 |||||  
 Db 25 DEPPQTPWDRVKDLATVYVEALKDSKDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELTYRQKVE 120  
 |||||  
 Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELTYRQKVE 144

## RESULT 3

JS0079

apolipoprotein A-I precursor - baboon

C:Species: Papio sp. (baboon)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 13-Jun-1997

C:Accession: JS0079

R:Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; Vandenberg, J.L.

Gene 74, 483-490, 1988

A:Title: The baboon gene for apolipoprotein A-I: characterization of a cDNA clone and

A:Reference number: JS0079; MUID:89232739

A:Accession: JS0079

A:Molecule type: mRNA

A:Residues: 1-267 &lt;HIX&gt;

A:Experimental source: liver

C:Comment: This protein is the principal protein component of high density lipoprotei

C:Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase

C:Comment: This protein contains a region of repeated amino acids which form amphipat

C:Genetics:

A:Gene: apoA1

C:Superfamily: apolipoprotein A-I

C:Keywords: HDL; lipid binding; lipoprotein

F:1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-267/Product: apolipoprotein A-I #status predicted &lt;LA1&gt;

F:123-144, 145-166, 167-188, 189-210, 211-232, 233-254/Region: tandem repeats

Query Match 96.0%; Score 606; DB 2; Length 267;  
 Best Local Similarity 95.0%; Pred. NO. 4.7e-40;  
 Matches 114; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
 |||||  
 Db 25 DEPPQTPWDRVKDLATVYVEALKDSKDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELTYRQKVE 120  
 |||||  
 Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELTYRQKVE 144

## RESULT 4

LPPDGA1

apolipoprotein A-I precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 17-Dec-1982 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C:Accession: A60940; A03092; A61418

R:Luo, C.C.; Li, W.H.; Chan, L.

J. Lipid Res. 30, 1735-1746, 1989

A:Title: Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implic

A:Reference number: A60940; MUID:90132271

A:Accession: A60940

A:Molecule type: mRNA

A:Residues: 1-266 &lt;LUO&gt;

R:Chung, H.; Randolph, A.; Reardon, I.; Heintz, R.L.

J. Biol. Chem. 257, 2961-2967, 1982

A:Title: The covalent structure of apolipoprotein A-I from canine high density lipopr

A:Reference number: A03092; MUID:82142425

A:Accession: A03092

A:Molecule type: protein

A:Residues: 25-167, 'G', 169-201, 'Q', 203-234, 'Q', 236-266 &lt;CHU&gt;

R:Nakai, T.; Whayne, T.F.; Tang, J.

FEBS Lett. 64, 409-411, 1976

A:Title: The amino- and carboxyl-terminal sequences of canine apolipoprotein A-I.

A:Reference number: A61418; MUID:76210910  
 A:Accession: A61418



A:Molecule type: protein  
A:Residues: 25-56,'Z','261-262','A' <NAK>  
C:Superfamily: apolipoprotein A-I  
C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid transport; 14  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 88.1%; Score 556; DB 1; Length 266;  
Best Local Similarity 88.1%; Pred. No. 3.6e-36;  
Matches 104; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDIATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSKLR 61  
: |||||  
Db 25 DDPQSPWDRVKDIATVYVDVLDKSGRDYVAQFEASALGKQLNLKLDNWDVSTSTFSKLR 84  
QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 119  
: |||||  
Db 85 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 142

## RESULT 5

apolipoprotein A-I - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
C:Accession: S31394  
R:Moehel, B.; Flach, R.; Weiss, B.; Weiler-Guetler, H.; Frey, A.; Zinke, H.; Gassen, H.  
submitted to the EMBL Data Library, November 1992  
A:Description: Genomic organization of the porcine apolipoprotein A1 gene and study of  
A:Reference number: S31394  
A:Accession: S31394  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-264 <MOE>  
A:Cross-references: EMBL:X69477; NID:g1889; PIDN:CAA49234.1; PID:g1890  
C:Superfamily: apolipoprotein A-I

Query Match 86.4%; Score 545; DB 2; Length 264;  
Best Local Similarity 85.6%; Pred. No. 2.5e-35;  
Matches 101; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDIATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSKLR 61  
: |||||  
Db 25 DDPQSPWDRVKDIATVYVDVLDKSGRDYVAQFEASALGKQLNLKLDNWDVSTSTFSKLR 84  
QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 119  
: |||||  
Db 85 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 142

## RESULT 6

apolipoprotein A-I - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
C:Accession: A46018  
R:Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.  
Genomics 15, 643-652, 1993  
A:Title: Characterization of the apolipoprotein A1 and CIII genes in the domestic pig.  
A:Reference number: A46018; MUID:93224154  
A:Accession: A46018  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-265 <BIR>  
A:Cross-references: GB:L00626; NID:g164358; PIDN:AAA30992.1; PID:g164359  
A:Note: sequence extracted from NCBI backbone (NCBIN:129509, NCBIPI:129511)  
C:Superfamily: apolipoprotein A-I

Query Match 86.4%; Score 545; DB 2; Length 265;

Best Local Similarity 85.6%; Pred. No. 2.5e-35;  
Matches 101; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDIATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSKLR 61  
: |||||  
Db 25 DDPQSPWDRVKDIATVYVDVLDKSGRDYVAQFEASALGKQLNLKLDNWDVSTSTFSKLR 84  
QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 119  
: |||||  
Db 85 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 142

## RESULT 7

apolipoprotein A-I - pig  
N:Alternate names: apo-A-I  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 11-May-2000  
C:Accession: JT0672; PN0471; A05311  
R:Trieu, V.N.; Patel, B.; Zhan, R.; Black, D.D.  
Gene 134, 267-270, 1993  
A:Title: Sequence of the porcine apoA-I gene.  
A:Reference number: JT0672; MUID:94085789  
A:Accession: JT0672  
A:Molecule type: DNA  
A:Residues: 1-265 <TRI>  
A:Cross-references: EMBL:Z14124; NID:g1893  
A:Note: this translation is not annotated in GenBank entry SSAPOAIG, release 111.0; t  
R:Trieu, V.N.; Hasler-Rapacz, J.; Rapacz, J.; Black, D.D.  
Gene 123, 173-179, 1993  
A:Title: Sequences and expression of the porcine apolipoprotein A-I and C-III mRNAs.  
A:Reference number: PN0471; MUID:93154581  
A:Accession: PN0471  
A:Molecule type: mRNA  
A:Residues: 105-265 <TR2>  
A:Experimental source: liver  
R:Mahley, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.  
Biochemistry 15, 1928-1933, 1976  
A:Title: Characterization of the plasma lipoproteins and apoproteins of the Erythroce  
A:Reference number: A90395; MUID:76184721  
A:Accession: A05311  
A:Molecule type: protein  
A:Residues: 25-34 <MAH>  
C:Comment: This protein is the major apolipoprotein of high-density lipoprotein and s  
C:Genetics:  
A:Gene: apoA-I  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid metabolism;  
F:99/Region: ochre stop codon

Query Match 84.5%; Score 533; DB 2; Length 265;  
Best Local Similarity 83.9%; Pred. No. 2.2e-34;  
Matches 99; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDIATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSKLR 61  
: |||||  
Db 25 DDPQSPWDRVKDIATVYVDVLDKSGRDYVAQFEASALGKQLNLKLDNWDVSTSTFSKLR 84  
QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 119  
: |||||  
Db 85 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKQWQEMETLYRQKV 142

## RESULT 8

apolipoprotein A-I precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 05-Jan-1996 #sequence\_revision 23-Aug-1997 #text\_change 13-Aug-1999  
C:Accession: I45853; A56858; A34649  
R:O'Huigin, C.; Chan, L.; Li, W.  
Mol. Biol. Evol. 7, 327-339, 1990

A:Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolution  
A:Reference number: I45853; MUID:90348478  
A:Accession: I45853  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-265 <OXH>  
A:Cross-references: GB:M35870; NID:g162677; PIDN:AAA30381.1; PID:g162678  
R:Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Auboiron, S.; Bauchart, D.; Chapman, M.J.; Gc Biochim. Biophys. Acta 1123, 145-150, 1992  
A:Title: Plasma lipid transport in the preterm infant calf, Bos spp: primary structure of h  
A:Reference number: A56858; MUID:92153895  
A:Accession: A56858  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 19-184, 'OL', 187-265 <SPA>  
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
A:Note: sequence extracted from NCBI backbone (NCBIF:83520)  
R:Auboiron, S.; Sparrow, D.A.; Beaubatie, L.; Bauchart, D.; Sparrow, J.T.; Laplaud, P.M. Biochem. Biophys. Res. Commun. 166, 833-839, 1990  
A:Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasma H  
A:Reference number: A34649; MUID:90147795  
A:Accession: A34649  
A:Molecule type: protein  
A:Residues: 25-70 <AUB>  
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
C:Superfamily: apolipoprotein A-I  
C:Keywords: lipid binding; lipoprotein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 82.6%; Score 521; DB 2; Length 265;  
Best Local Similarity 81.4%; Pred. No. 1.9e-33;  
Matches 96; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLDNWDSTVSTFSKLR 61  
: ||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| : |||||  
Db 25 DDPQSSWDRVKDFATVYVEAIKDSGRDYVAQFEASALGKQLNLKLDNWDSTVSTFSKVR 84  
QY 62 EQLGPVTQEFWMDLKEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELRYQKV 119  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||||  
Db 85 EQLGPVTQEFWMDLKEKETASLRQEMHKDLLEEVKQKVQPYLDLDFQKKWHEVEIYRQKV 142

RESULT 9  
LPRB12  
apolipoprotein A-I precursor (clone 2Zap AI) - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S06064  
R:Paraskevopoulou, T.B.; Kritis, A.; Zannis, V.  
submitted to the EMBL Data Library, July 1989  
A:Reference number: S06064  
A:Accession: S06064  
A:Molecule type: mRNA  
A:Residues: 1-266 <PAR>  
A:Cross-references: EMBL:X15908; NID:g1457; PIDN:CAA34024.1; PID:g1458  
C:Comment: This protein is synthesized in the small intestine.  
C:Comment: This protein is a major component of the high density lipoproteins in plasma.  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe  
F:1-18/Domain: signal sequence #status predicted <PRO>  
F:19-24/Domain: propeptide #status predicted <SIG>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 82.6%; Score 521; DB 1; Length 266;  
Best Local Similarity 79.0%; Pred. No. 1.9e-33;  
Matches 94; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLDNWDSTVSTFSKLR 61  
: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||||  
Db 25 DEPRSSWMDIKDFATVYVDVTKDSGREYVAQFEASAFGKQLNLKLDNWDSTVSTFSKLR 84

QY 62 EQLGPVTQEFWMDLKEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELRYQKV 120  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||||  
Db 85 EQLGPVTQEFWMDLKEKETEGLRQEMSKDLQEVKQKVQPYLDLDFQKKWQEMELRYQKV 143

RESULT 10  
LPRB1B  
apolipoprotein A-I precursor (clone PRBA-502) - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
R:Pan, T.C.; Hao, Q.L.; Yamin, T.T.; Dai, P.H.; Chen, B.S.; Chen, S.L.; Kroon, P.A.; Eur. J. Biochem. 170, 99-104, 1987  
A:Title: Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit apolipoprotein  
A:Reference number: S00230; MUID:88082866  
A:Accession: S00230  
A:Molecule type: mRNA  
A:Residues: 1-265 <PAN>  
A:Cross-references: EMBL:X06658; NID:g1461; PIDN:CAA29857.1; PID:g1462  
A:Note: the authors translated the codon AGC for residue 174 as Arg  
A:Accession: S20557  
A:Molecule type: DNA  
A:Residues: 1-17, 'R', 19-44, 'I', 46-122, 'Y', 124-146, 'V', 148-265 <PAN2>  
A:Cross-references: EMBL:X06659; NID:g1459; PIDN:CAA29858.1; PID:g1460  
C:Comment: This protein is synthesized in the small intestine.  
C:Comment: This protein is a major component of the high density lipoproteins in plas  
C:Genetics:  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 81.9%; Score 517; DB 1; Length 265;  
Best Local Similarity 78.2%; Pred. No. 3.8e-33;  
Matches 93; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLDNWDSTVSTFSKLR 61  
: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||||  
Db 25 DEPRSSWMDIKDFATVYVDVTKDSGREYVAQFEASAFGKQLNLKLDNWDSTVSTFSKLR 84  
QY 62 EQLGPVTQEFWMDLKEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELRYQKV 120  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||||  
Db 85 EQLGPVTQEFWMDLKEKETEGLRQEMSKDLQEVKQKVQPYLDLDFQKKWQEMELRYQKV 143

RESULT 11  
A24998  
apolipoprotein A-I - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 31-Dec-1993  
C:Accession: A24998  
R:Yang, C.; Yang, T.; Pownall, H.J.; Gotto Jr., A.M.  
Eur. J. Biochem. 160, 427-431, 1986  
A:Title: The primary structure of apolipoprotein A-I from rabbit high-density lipopro  
A:Reference number: A24998; MUID:87030294  
A:Accession: A24998  
A:Molecule type: protein  
A:Residues: 1-241 <YAN>  
C:Superfamily: apolipoprotein A-I  
C:Keywords: HDL; lipid binding; lipoprotein

Query Match 81.9%; Score 516.5; DB 2; Length 241;  
Best Local Similarity 80.7%; Pred. No. 3.7e-33;  
Matches 96; Conservative 14; Mismatches 8; Indels 1; Gaps 1;

QY 2 EPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLDNWDSTVSTFSKLR 61  
: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||||  
Db 1 DEPRSSWMDIKDFATVYVDV-KDSGREYVAQFEASAFGKQLNLKLDNWDSTVSTFSKLR 59



A:Accession: A05314  
A:Molecule type: protein  
A:Residues: 1-18,'X',20-21,'X',23-28,'X',30-38,'X',40-41,'X',43-45 <GOR>  
C:Comment: This protein is synthesized in the liver and small intestine. The propeptide  
C:Comment: This protein is a major component of the high density lipoproteins in plasma.  
C:Genetics:  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lipid  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-259/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 75.9%; Score 479; DB 2; Length 259;  
Best Local Similarity 73.9%; Pred. No. 3.3e-30;  
Matches 88; Conservative 14; Mismatches 17; Indels 0; Gaps 0;  
QY 2 EPPQSPWDRYKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTFSKLR 61  
: ||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| : || : ||  
Db 25 DEPQSQWDRYKDFATVYVDVDAVKDSGRDYVSQFESSLGLKQLNLNLNDNWDTLGSTVGRLQ 84  
QY 62 EQLGPTYQEEFNDLEKETEGLRQEMSKDLEEVAKAKVQPYLDDFQKKWQEEMLYRQKVE 120  
: ||||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| : || : ||  
Db 85 EQLGPTYQEEFWANLEKETDNLRNEMNKDLNVKQKMPHLDEFQEKWNEVEAYRQKLE 143

Search completed: September 22, 2002, 12:06:29  
Job time: 324 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:22:17 ; Search time 35.02 Seconds  
(without alignments)  
132.677 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_144

Perfect score: 631

Sequence: 1 DEPPQSPWDRVKDLATVVD.....LDDFOKKWQEMELYRQKE 120

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	631	100.0	267	1 APAL_HUMAN	P02647 homo sapien
2	606	96.0	267	1 APAL_MACRA	P15568 macaca fasc
3	556	88.1	266	1 APAL_CANFA	P02648 canis faml
4	545	86.4	265	1 APAL_PIG	P18648 sus scrofa
5	521	82.6	265	1 APAL_BOVIN	P15497 bos taurus
6	521	82.6	266	1 APAL_RABIT	P09809 oryctolagus
7	481	76.2	264	1 APAL_MOUSE	Q00623 mus musculu
8	479	75.9	259	1 APAL_RAT	P04639 rattus norv
9	463	73.4	265	1 APAL_TUPGB	O18759 tupaia glis
10	329	52.1	264	1 APAL_CHICK	P08250 gallus gall
11	328	52.0	264	1 APAL_COTTA	P32918 coturnix co
12	326	51.7	264	1 APAL_ANAPL	O42296 anas platyr
13	142	22.5	429	1 APAL_MACFA	P33621 macaca fasc
14	139	22.0	396	1 APAL_HUMAN	P06727 homo sapien
15	138	21.9	401	1 APAL_PAPAN	Q28758 papio anubi
16	125.5	19.9	391	1 APAL_RAT	P02651 rattus norv
17	121	19.2	382	1 APAL_PIG	O46409 sus scrofa
18	114.5	18.1	262	1 APAL_ONCMY	O57523 oncorhynch
19	114	18.1	395	1 APAL_MOUSE	P06728 mus musculu
20	112.5	17.8	262	1 APAL_BRARE	O42363 brachydant
21	108	17.1	281	1 APE_BRARE	O42364 brachydant
22	107	17.0	262	1 AP12_ONCMY	O57524 oncorhynch
23	106.5	16.9	262	1 APAL_SALTR	O91488 salmo trutt
24	104	16.5	312	1 APE_RAT	P02650 rattus norv
25	101.5	16.1	365	1 RECF_CHLTR	O84077 chlamydia t
26	101	16.0	20	1 APAL_ERYP	P18647 erythrocebu
27	98	15.5	298	1 APE_CAVPO	P23529 cavia porce
28	98	15.5	311	1 APE_MOUSE	P08226 mus musculu
29	95.5	15.1	260	1 APAL_SPAU	O42175 sparus aura
30	90	14.3	317	1 APE_MACFA	P10517 macaca fasc
31	90	14.3	317	1 APE_PAPAN	P05770 papio anubi
32	89	14.1	305	1 APE_CANFA	P18649 canis faml
33	88.5	14.0	539	1 MYS3_HYDAT	P39922 hydra atten

34	87	13.8	317	1 APE_PIG	P18650 sus scrofa
35	85	13.5	107	1 APE_MACMU	Q28502 macaca mula
36	85	13.5	258	1 APAL_SALSA	P27007 salmo salar
37	85	13.5	657	1 LAMA_CHICK	P13648 gallus gall
38	84	13.3	583	1 LAM1_XENLA	P09010 xenopus lae
39	83	13.2	107	1 APE_SAISC	Q28995 saimiri sci
40	83	13.2	263	1 CENE_HUMAN	Q02224 homo sapien
41	82.5	13.1	365	1 RECF_CHLMU	Q9PKW5 chlamydia m
42	82.5	13.1	512	1 K2C5_XENLA	P16878 xenopus lae
43	82.5	13.1	574	1 LAMC_MOUSE	P15156 mus musculu
44	82.5	13.1	664	1 LAMA_HUMAN	P02545 homo sapien
45	82.5	13.1	665	1 LAMA_MOUSE	P48678 mus musculu

#### ALIGNMENTS

RESULT	ID	APAL_HUMAN	STANDARD	PRT	267 AA.
AC	P02647				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).				
GN	APOA1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=84221405; PubMed=6328445;				
RX	Medline=84221405; PubMed=6328445;				
RA	Seilhamer J.J., protter A.A., Frossard P., Levy-Wilson B.;				
RA	Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,				
RA	Baralle F.E.;				
RT	"Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA				
RT	abundance.";				
RL	Nucleic Acids Res. 12:3917-3932(1984).				
RN	[2]	SEQUENCE FROM N.A.			
RP	MEDLINE=8502665; PubMed=6207999;				
RX	Medline=8502665; PubMed=6207999;				
RA	Seilhamer J.J., protter A.A., Frossard P., Levy-Wilson B.;				
RA	Shoulders C.C., Kornblitt A.R., Munro B.S., Baralle F.E.;				
RT	"Gene structure of human apolipoprotein AI.";				
RL	Nucleic Acids Res. 11:2827-2837(1983).				
RN	[3]	SEQUENCE FROM N.A.			
RP	MEDLINE=83220822; PubMed=6406984;				
RX	Medline=83220822; PubMed=6406984;				
RA	Shoulders C.C., Kornblitt A.R., Munro B.S., Baralle F.E.;				
RT	"Gene structure of human apolipoprotein AI.";				
RL	Nucleic Acids Res. 11:2827-2837(1983).				
RN	[4]	SEQUENCE FROM N.A.			
RP	MEDLINE=83220772; PubMed=6304641;				
RX	Medline=83220772; PubMed=6304641;				
RA	Cheung P., Chan L.;				
RT	"Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";				
RL	Nucleic Acids Res. 11:3703-3715(1983).				
RN	[5]	SEQUENCE FROM N.A.			
RP	MEDLINE=84119464; PubMed=6198645;				
RX	Medline=84119464; PubMed=6198645;				
RA	Law S.W., Brewer H.B. Jr.;				
RT	"Nucleotide sequence and the encoded amino acids of human				
RT	apolipoprotein A-I mRNA.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).				
RN	[6]	SEQUENCE FROM N.A.			
RP	MEDLINE=86008382; PubMed=2995392;				
RX	Medline=86008382; PubMed=2995392;				
RA	Law S.W., Brewer H.B. Jr.;				
RT	"Tangier disease. The complete mRNA sequence encoding for				
RT	preproapo-A-I.";				
RL	J. Biol. Chem. 260:12810-12814(1985).				
RN	[7]				

RP SEQUENCE FROM N.A.  
RX MEDLINE=84016011; PubMed=6413973;  
RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
RT "Isolation and characterization of the human apolipoprotein A-I  
RT gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89377481; PubMed=2673706;  
RA Mogilevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,  
RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
RA Holmquist L., Carlson L.A., Bollen A.;  
RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
RT coli: purification and biochemical characterization.";  
RL DNA 8:429-436(1989).  
RN [9]  
RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
RX MEDLINE=88196137; PubMed=3129297;  
RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
RA Zannis V.I.;  
RT "Sequence and expression of Tangier apoA-I gene.";  
RL Eur. J. Biochem. 173:465-471(1988).  
RN [10]  
RP SEQUENCE OF 118-267 FROM N.A.  
RX MEDLINE=83091059; PubMed=6294659;  
RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
RT "Isolation and characterization of cDNA clones for human  
RT apolipoprotein A-I.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
RN [11]  
RP SEQUENCE OF 19-27.  
RX MEDLINE=83256553; PubMed=6409108;  
RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
RA Light J.A.;  
RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
RN [12]  
RP SEQUENCE OF 25-267.  
RX MEDLINE=78123731; PubMed=204308;  
RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
RA Bronzert T.J.;  
RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated  
RT from high density lipoproteins.";  
RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
RN [13]  
RP SEQUENCE OF 25-267.  
RX MEDLINE=75133493; PubMed=164450;  
RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;  
RT "The primary structure of human plasma high density apolipoprotein  
RT glutamine I (ApoA-I). II. The amino acid sequence and alignment of  
RT cyanogen bromide fragments IV, III, and I.";  
RL J. Biol. Chem. 250:2725-2738(1975).  
RN [14]  
RP SEQUENCE OF 25-56.  
RX MEDLINE=88331387; PubMed=3047170;  
RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
RT A-I (Apo A-I). A novel function of Apo A-I.";  
RL J. Clin. Invest. 82:803-807(1988).  
RN [15]  
RP SEQUENCE OF 25-48.  
RX MEDLINE=89380318; PubMed=2506184;  
RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
RA Chappdelaine A.;  
RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
RT proteins.";  
RL J. Biol. Chem. 264:16853-16857(1989).  
RN [16]  
RP SEQUENCE OF 25-43.  
RX MEDLINE=88070603; PubMed=3120314;  
RA Prioli R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
RA Pereira M.E.A.;  
RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
RT neuraminidase, to high-density lipoprotein.";  
RL Science 238:1417-1419(1987).  
RN [17]  
RP SEQUENCE OF 25-42.  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [18]  
RP PALMITOYLATION.  
RX MEDLINE=86140194; PubMed=3005308;  
RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
RT "Human apolipoprotein A-I. Post-translational modification by fatty  
RT acid acylation.";  
RL J. Biol. Chem. 261:3911-3914(1986).  
RN [19]  
RP PROCESSING.  
RX MEDLINE=83195100; PubMed=6405383;  
RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
RA Breslow J.L.;  
RT "Intracellular and extracellular processing of human apolipoprotein  
RT A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
RN [20]  
RP STRUCTURE BY NMR OF 190-209.  
RX MEDLINE=96270776; PubMed=8664326;  
RA Wang G., Treleaven W.D., Cushley R.J.;  
RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
RT presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
RT and CD. Evidence for specific peptide-SDS interactions.";  
RL Biochim. Biophys. Acta 1301:174-184(1996).  
RN [21]  
RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
RX MEDLINE=98024124; PubMed=9356442;  
RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
RT lipid-bound conformation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
RN [22]  
RP VARIANT MILANO.  
RX MEDLINE=83109095; PubMed=6401735;  
RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,  
RA Franceschini G., Sirtori C.R.;  
RT "Apolipoprotein A-IMilano. Detection of normal A-I in affected  
RT subjects and evidence for a cysteine for arginine substitution in the  
RT variant A-I.";  
RL J. Biol. Chem. 258:2508-2513(1983).  
RN [23]  
RP VARIANT TANGIER.  
RX MEDLINE=83300108; PubMed=6412234;  
RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;  
RT "Tangier disease: defective recombination of a specific Tangier  
RT apolipoprotein A-I isoform (pro-apo A-I) with high density  
RT lipoproteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).  
RN [24]  
RP VARIANT NORWAY.  
RX MEDLINE=84289383; PubMed=6432779;  
RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,  
RA Utermann G., Haas J., Steinmetz A., Menzel H.J., Assmann G.;  
RT "Abnormal lecithin:cholesterol acyltransferase activation by a human  
RT apolipoprotein A-I variant in which a single lysine residue is  
RT deleted.";  
RL J. Biol. Chem. 259:10063-10070(1984).  
RN [25]  
RP SEQUENCE OF 25-107 (VARIANT IOWA).  
RX MEDLINE=89050104; PubMed=3142462;  
RA Nichols W.C., Dwulet F.E., Liepnieks J., Benson M.D.;  
RT "Variant apolipoprotein AI as a major constituent of a human  
RT hereditary amyloid.";

Query Match 100.0%; Score 631; DB 1; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-44;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 25 DEPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
 QY 61 REQLGPVTQEFWNLKEKTEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELLYRQKVE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 85 REQLGPVTQEFWNLKEKTEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELLYRQKVE 144

RESULT 2  
 APAL\_MACFA STANDARD; PRT; 267 AA.

AC P15568; P17929;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and  
 OS Papio hamadryas (Hamadryas baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OC NCBI\_TaxID=9541, 9557;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.fascicularis;  
 RX MEDLINE=87191989; PubMed=3106152;  
 RA Polites H.G., Melchior G.W., Castle C.K., Marotti K.R.;  
 RT "The primary structure of cynomolgus monkey apolipoprotein A-1  
 deduced from the cDNA sequence: comparison to the human sequence.";  
 RL Gene 49:103-110(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.fascicularis;  
 RX MEDLINE=92305062; PubMed=1610902;  
 RA Murray R.W., Marotti K.R.;  
 RT "Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene  
 and corresponding flanking regions.";  
 RL Biochim. Biophys. Acta 1131:207-210(1992).  
 RN [3]  
 RP SEQUENCE OF 25-48.  
 RC SPECIES=M.fascicularis;  
 RX MEDLINE=87185451; PubMed=3105581;  
 RA Herbert P.N., Bausserman L.L., Lynch K.M., Saritelli A.L.,  
 RA Kantor M.A., Nicolosi R.J., Shulman R.S.;  
 RT "Homologues of the human C and A apolipoproteins in the Macaca  
 fascicularis (cynomolgus) monkey.";  
 RL Biochemistry 26:1457-1463(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.hamadryas; TISSUE=Liver;  
 RX MEDLINE=89232739; PubMed=2907746;  
 RA Hixson J.E., Borenstein S., Cox L.A., Rainwater D.L., Vandeberg J.L.;  
 RT "The baboon gene for apolipoprotein A-I: characterization of a cDNA  
 clone and identification of DNA polymorphisms for genetic studies of  
 cholesterol metabolism.";  
 RL Gene 74:483-490(1988).  
 RN [5]  
 RP SEQUENCE OF 1-10 FROM N.A.  
 RC SPECIES=M.fascicularis;  
 RA Sorci-Thomas M.;  
 RL Submitted (Oct-1991) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCEPTION BY PROMOTING  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.

CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M15411; AAA36834.1; -  
 DR EMBL; M83242; AAA36832.1; -  
 DR EMBL; M35634; AAA35380.1; -  
 DR EMBL; M69223; AAA36831.1; -  
 DR PIR; A26529; A26529.  
 DR PIR; A26627; A26627.  
 DR PIR; JS0079; JS0079.  
 DR PIR; S23135; S23135.  
 DR HSSP; P02647; 10DR.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 1.  
 DR Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 KW SIGNAL  
 FT 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 267  
 FT DOMAIN 68 267  
 FT REPEAT 68 89  
 FT REPEAT 90 111  
 FT REPEAT 112 122  
 FT REPEAT 123 144  
 FT REPEAT 145 166  
 FT REPEAT 167 188  
 FT REPEAT 189 210  
 FT REPEAT 211 232  
 FT REPEAT 233 243  
 FT REPEAT 244 267  
 FT CONFLICT 13 13 L -> P (IN REF. 1).  
 SQ SEQUENCE 267 AA; 30735 MW; 869955C024088E21 CRC64;

Query Match 96.0%; Score 606; DB 1; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 5.3e-42;  
 Matches 114; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 25 DEPPQSPWDRVKDLATVYVDVLDKSGRDVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
 QY 61 REQLGPVTQEFWNLKEKTEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELLYRQKVE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 85 REQLGPVTQEFWNLKEKTEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELLYRQKVE 144

RESULT 3

APAL\_CANFA STANDARD; PRT; 266 AA.  
 AC P02648;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=90132271; PubMed=2515239;  
 RA Luo C.-C., Li W.-H., Chan L.;

```

RT "Structure and expression of dog apolipoprotein A-I, E, and C-I
RT mRNAs: implications for the evolution and functional constraints of
RT apolipoprotein structure."
RL J. Lipid Res. 30:1735-1746(1989).
RN [2]
RX MEDLINE=82142425; PubMed=6801039;
RA Chung H., Randolph A., Reardon I., Heinrichson R.L.;
RT "The covalent structure of apolipoprotein A-I from canine high
RT density lipoproteins."
RL J. Biol. Chem. 257:2961-2967(1982).
RN [3]
RP SEQUENCE OF 25-57 AND 262-265.
RX MEDLINE=76210910; PubMed=179887;
RA Nakai T., Whayne T.F., Tang J.;
RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein
RT A-I."
RL FEBS Lett. 64:409-411(1976).
RN [4]
RP SEQUENCE OF 25-37.
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins."
RL Electrophoresis 18:2795-2802(1997).
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS. SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
DR HSP; P02647; 10DR.
DR HSC-2DPAGE; P02648; DOG.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; signal.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24
FT CHAIN 25 266
FT DOMAIN 67 266
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 266
FT CONFLICT 168 168 A->G (IN REF. 2).
FT CONFLICT 202 202 E->Q (IN REF. 2).
FT CONFLICT 235 235 E->Q (IN REF. 2).
FT CONFLICT 264 266 NAO->A (IN REF. 3).
SQ SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;

Query Match 88.1%; Score 556; DB 1; Length 266;
Best Local Similarity 88.1%; Pred. No. 5.4e-38;
Matches 104; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 2 EPPQSPMDRVKDLATVYVDLKDGRDYSQFEGSALGKOLNKLIDNWDSTSTFSKLR 61
Db : ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 25 DEQSPMDRVKDLATVYVDLKDGRDYSQFEGSALGKOLNKLIDNWDSTSTFSKLR 84
OY 62 EQIGPYTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQKV 119
Db 85 EQIGPYTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQKV 142

```

```

RESULT 4
ID APAL_PIG STANDARD; PRT; 265 AA.
AC P18648;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93224154; PubMed=8468059;
RA Birchbauer A., Knipping G., Juritsch B., Aschauer H., Zechner R.;
RT "Characterization of the apolipoprotein AI and CIII genes in the
RT domestic pig."
RL Genomics 15:643-652(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94125128; PubMed=8294940;
RA Moeckel B., Zinke H., Flach R., Weiss B., Weiler-Guettler H.,
RT "Expression of apolipoprotein A-I in porcine brain endothelium in
RT vitro."
RL J. Neurochem. 62:788-798(1994).
RN [3]
RP SEQUENCE OF 34-265 FROM N.A.
RX MEDLINE=90132667; PubMed=2105375;
RA Weiler-Guettler H., Sommerfeldt M., Papandrikopoulou A., Mischek U.,
RT Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;
RT "Synthesis of apolipoprotein A-I in pig brain microvascular
RT endothelial cells."
RL J. Neurochem. 54:444-450(1990).
RN [4]
RP SEQUENCE OF 105-265 FROM N.A.
RX MEDLINE=93154581; PubMed=8428656;
RA Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;
RT "Sequences and expression of the porcine apolipoprotein A-I and C-III
RT mRNAs."
RL Gene 123:173-179(1993).
RN [5]
RP SEQUENCE OF 25-265.
RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,
RL Jackson K., Gustavsson I., Rapacz J.;
RN Submitted (OCT-1995) to the SWISS-PROT data bank.
RN [6]
RP SEQUENCE OF 25-34.
RX MEDLINE=76184721; PubMed=178359;
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
RT "Characterization of the plasma lipoproteins and apoproteins of the
RT Erythrocebus patas monkey."
RL Biochemistry 15:1928-1933(1976).
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE
CC LIVER.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```





```
RESULT 6
ID APAL_RABIT STANDARD; PRT; 266 AA.
AC P09809;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22AP AI; TISSUE=Small intestine;
RA Paraskevopoulou T.B., Kritis A., Zannis V.I.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=88082866; PubMed=3121329;
RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,
RT Kroon P.A., Chao Y.S.;
RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit
RT apolipoprotein A-I is synthesized in the intestine but not in the
RT liver.";
RL Eur. J. Biochem. 170:99-104(1987).
RN [3]
RP SEQUENCE OF 25-266.
RX MEDLINE=87030294; PubMed=3095115;
RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;
RT "The primary structure of apolipoprotein A-I from rabbit high-density
RT lipoprotein.";
RL Eur. J. Biochem. 160:427-431(1986).
RN [4]
RP FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC [5]
CC SUBCELLULAR LOCATION: Extracellular.
CC [6]
CC TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC [7]
CC SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15908; CAA34024.1; -
DR EMBL: X06658; CAA29857.1; -
DR EMBL: X06659; CAA29858.1; -
DR PIR: S06064; LPRB12.
DR PIR: S00230; LPRB1B.
DR PIR: A24998; A24998.
DR HSSP: P02647; 1AV1.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein.1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT STIGMAL 1 18
FT PROPEP 19 24
FT CHAIN 25 266
FT DOMAIN 67 266
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
```

```
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 266
FT CONFLICT 18 18
FT CONFLICT 44 44
FT CONFLICT 45 45
FT CONFLICT 107 107
FT CONFLICT 123 123
FT CONFLICT 147 147
FT CONFLICT 150 150
FT CONFLICT 191 191
FT CONFLICT 195 195
FT CONFLICT 211 211
FT CONFLICT 255 256
FT CONFLICT 256 256
SQ SEQUENCE 266 AA; 30591 MW; 0FF6DB386497C7D2 CRC64;
8.
9 (HALF-LENGTH).
10.
A -> R (IN REF. 2; CAA29858).
MISSING (IN REF. 3).
V -> I (IN REF. 2; CAA29858).
E -> Q (IN REF. 3).
Y -> F (IN REF. 2; CAA29857).
A -> V (IN REF. 2; CAA29858 AND REF. 3).
R -> G (IN REF. 2; CAA29858 AND
CAA29857).
N -> O (IN REF. 3).
MISSING (IN REF. 2; CAA29858 AND
CAA29857).
S -> K (IN REF. 2; CAA29858 AND
CAA29857).
VL -> LV (IN REF. 3).
L -> V (IN REF. 2; CAA29858 AND
CAA29857).
```

Query Match 82.6%; Score 521; DB 1; Length 266;  
Best Local Similarity 79.0%; Pred. No. 3.5e-35;  
Matches 94; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

```
QY 2 EPPQSPMDRVKDLATVYVDVLKDSGRDVSOFEGSALGKQLNKLDDNWDSTVFSKLR 61
Db 25 DEPRSSWDKIKDFATVYVDVYKDSGREYVAQFEASAFGKQLNKLDDNWDSTVSKLQ 84
QY 62 EQLGPVTOEFWDNLEKETEGLEQEMSKDLEEVKAKQVPLDDFQKKWQEMELYRQKE 120
Db 85 EQLGPVTOEFWDNLEKETEGLEQEMSKDLEEVKAKQVPLDDFQKKWQEMELYRQKE 143
```

```
RESULT 7
ID APAL_MOUSE STANDARD; PRT; 264 AA.
AC Q00623;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92281682; PubMed=1596360;
RA Stoffel W., Mueller R., Binczek E., Hofmann K.;
RT "Mouse apolipoprotein AI. cDNA-derived primary structure, gene
RT organisation and complete nucleotide sequence.";
RL Biol. Chem. Hoppe-Seyler 373:187-193(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93122774; PubMed=1478650;
RA Janusz J.L., Azrolan N., O'Connell A., Aalto-Setälä K., Breslow J.L.;
RT "Characterization of the mouse apolipoprotein ApoA-1/Apoc-3 gene
RT locus: genomic, mRNA, and protein sequences with comparisons to other
RT species.";
RL Genomics 14:1081-1088(1992).
RN [3]
RP FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC [4]
CC SUBCELLULAR LOCATION: Extracellular.
CC [5]
CC TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC [6]
CC SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; X64262;	CAA45560.1;	-.
DR	EMBL; X64263;	CAA45561.1;	-.
DR	EMBL; L04149;	-;	NOT_ANNOTATED_CDS.
DR	EMBL; L04151;	-;	NOT_ANNOTATED_CDS.
DR	PIR; S22420;	S22420.	
DR	PIR; A44364;	A44364.	
DR	HSSP; P02647;	IAVI.	
DR	SWISS-2DPAGE; Q00623;	MOUSE.	
DR	MGI; MGI:88049;	Apoal.	
DR	InterPro; IPR000074;	Apolipoprotein.	
DR	Pfam; PF01442;	Apolipoprotein; 1.	
KW	plasma; Lipid transport;	HDL; Cholesterol metabolism;	Repeat; Signal..
FT	SIGNAL	1	18
FT	PROPEP	19	24
FT	CHAIN	25	264
FT	DOMAIN	67	264
FT	REPEAT	67	88
FT	REPEAT	89	110
FT	REPEAT	111	121
FT	REPEAT	122	143
FT	REPEAT	144	165
FT	REPEAT	166	187
FT	REPEAT	188	207
FT	REPEAT	208	229
FT	REPEAT	230	240
FT	REPEAT	241	264
FT	SEQUENCE	264 AA;	30587 MW; C453FF2019634AAC CRC64;

Query Match	76.28;	Score 481;	DB 1;	length 264;
Best Local Similarity	72.9%;	Pred. No. 5.6e-32;		
Matches	86;	Conservative	21;	Mismatches 11;
				Indels 0;
				Gaps 0;

[illegible]

RESULT	8
APAL_RAT	
ID	APAL_RAT STANDARD; PRT; 259 AA.
AC	P04639;
DT	13-AUG-1987 (Rel. 05, Created)
DT	13-AUG-1987 (Rel. 05, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Apolipoprotein A-I precursor (Apo-AI).
GN	APoA1.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=84207987; PubMed=6426956;
RA	Poncin J.E., Martial J.A., Gjelten J.E.;
RT	"Cloning and structure analysis of the rat apolipoprotein A-I cDNA.";
RL	Eur. J. Biochem. 140:493-498(1984).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=87008540; PubMed=3020028;
RA	Haddad I.A., Ordovas J.M., Fitzpatrick T., Karathanasis S.K.;
RT	"Linkage, evolution, and expression of the rat apolipoprotein A-I, C-

RT III, and A-IV genes.<sup>11</sup>  
 RL J. Biol. Chem. 261:13268-13277(1986).  
 RN [3]  
 RP SEQUENCE OF 1-45.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=82098162; PubMed=6798036;  
 RA Gordon J.I., Smith D.P., Andy R., Alpers D.H., Schonfeld G.,

"the primary translation product of rat intestinal apolipoprotein A-I  
 mRNA is an unusual preproprotein.";  
 J. Biol. Chem. 257:971-978(1982).  
 -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 -I- SUBCELLULAR LOCATION: Extracellular.  
 -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CHYLOMICRONS.  
 -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; M000001; AAAA0749.1; -.
DR	EMBL; X00558; CAA25224.1; -.
DR	EMBL; J02597; AAA40745.1; -.
DR	PIR; A24700; A24700.
DR	HSSP; P02647; LAVI.
DR	InterPro; IPR000074; Apolipoprotein.
KW	Pfam; PF01442; Apolipoprotein; 1.
DR	Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal
FT	SIGNAL
FT	PROPEP
FT	CHAIN
FT	DOMAIN
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	CONFLICT
FT	CONFLICT
FT	CONFLICT
SEQ	SEQUENCE

Query Match	75.98;	Score 479;	DB 1;	Length 259;
Best Local Similarity	73.98;	Pred. No. 7.9e-32;		
Matches 88;	Conservative 14;	Mismatches 17;	Indels 0;	Gaps 0;

[illegible]

RESULT	9		
APAI_TUPGB			
ID	APAI_TUPGB	STANDARD;	PRT;
AC	018759;		265 AA.

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF005638; AAB82326.1; -.
DR HSSP; P02647; 1AV1.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 265
FT DOMAIN 67 265
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 265
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAFA165 CRC64;

```

```

Query Match 73.4%; Score 463; DB 1; Length 265;
Best Local Similarity 69.7%; Pred. No. 1.6e-30;
Matches 83; Conservative 19; Mismatches 17; Indels 0; Gaps 0;

```

```

QY 2 EPPQSPWDRYKDLATVYVDVLKDSGRDVSQFEGSALGKQLNKLDDNWDVSFSPSKLR 61
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 DEPOSSWDVRVRLANVYVDVAVKESGREYVSQLASALGKQLNKLVDNWDTLGTFQKVH 84
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 EQLGPTQGFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 EHLGPAQGFWEKLEKETELRRLINKDLEDVROKTPFLDEIQKKWQEDLERYRQKVE 143
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
APAL_CHICK STANDARD; PRT; 264 AA.
AC P08250;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88049703; PubMed=3118875;
RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.;
RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and
RT evolution."
RL Biochem. Biophys. Res. Commun. 148:485-492(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88152500; PubMed=3126099;
RA Ferrari S., Tarugi P., Drusiani E., Calandra S., Fregni M.;
RT "The complete sequence of chick apolipoprotein AI mRNA and its
RT expression in the developing chick."
RL Gene 60:39-46(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87222301; PubMed=3108248;
RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,
RA Leberer H., Lusis A.J.;
RT "Structure, evolution, and regulation of chicken apolipoprotein A-I."
RL J. Biol. Chem. 262:7058-7065(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92381402; PubMed=1512510;
RA Lamou-Fava S., Sastry R., Ferrari S., Rajavashisth T.B.,
RA Lusis A.J., Karathanasis S.K.;
RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene
RT expression: differences between avian and mammalian apoA-I gene
RT transcription control regions."
RL J. Lipid Res. 33:831-842(1992).
RN [5]
RP SEQUENCE OF 25-44.
RX MEDLINE=83213468; PubMed=6406496;
RA Shackelford J.E., Leberer H.G.;
RT "Synthesis and secretion of apolipoprotein AI by chick breast
RT muscle."
RL J. Biol. Chem. 258:7175-7180(1983).
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M17961; AAA48593.1; -.
DR EMBL; M18746; AAA48594.1; -.
DR EMBL; M25559; AAA48592.1; -.
DR EMBL; M96012; AAA48597.1; -.
DR PIR; S01453; LPECH1.
DR PIR; JH0471; JH0471.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24

```



FT	CHAIN	25	264	APOLIPROTEIN A-I.
FT	DOMAIN	67	264	10 X APPROXIMATE TANDEM REPEATS
FT	REPEAT	67	88	1.
FT	REPEAT	89	110	2.
FT	REPEAT	111	121	3 (HALF-LENGTH).
FT	REPEAT	122	143	4.
FT	REPEAT	144	165	5.
FT	REPEAT	166	187	6.
FT	REPEAT	188	209	7.
FT	REPEAT	210	231	8.
FT	REPEAT	232	242	9 (HALF-LENGTH).
FT	REPEAT	243	264	10.
FT	CONFLICT	16	16	T -> I (IN REF. 4).
FT	CONFLICT	148	148	E -> K (IN REF. 3).
SO	SEQUENCE	264 AA;	30680 MW;	968320E81E2AC5C2 CRC64;

Query Match	52.1%;	Score 329;	DB 1;	Length 264;
Best Local Similarity	50.0%;	Pred. No. 8.9e-20;		
Matches 59;	Conservative 30;	Mismatches 29;	Indels 0;	Gaps 0;

```
QY      2 EPOS PWD RVK D L A T V Y V D V L K D S G R D Y V S O F E G S A L G K O I N L K L D N W D S V T S T E S K I R    61
       : | : | | : : | : : : | : | : | : | : | : | : | : : : | : |
Db     25 DEOP Q T P L D R I R D M V D V L E T V K A S G K D A I A O F E S S A V G K Q I D L K L A D N L D T L S A A A A K I R    84

QY      62 EQLG P V T Q E F W D N L E K E T E G L R O E M S K D L E E V K A K V Q P Y L D D E Q K K W O E M E L Y R O K V    119
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db     85 EDMA P Y K E V R E M W L K D T E A L R A E L T K D L E E V K E K I R P F L D O F S A K W T E I L E Q Y R O L    142
```

RESULT	11	
APAL_COTJA		
ID	APAL_COTJA	STANDARD; PRT; 264 AA

DT 01-OCT-1993 (Rel. 27, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Coturnix coturnix japonica (Japanese quail).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
OX NCBI\_TaxID=93934;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=97212027; PubMed=9058967;  
RA Oku H., Toda T., Nagata J., Ishikawa M., Neyazaki K., Shinjyo C.,  
RA Chinen I.;  
RA "Apolipoprotein A-1 of Japanese quail: cDNA sequence and modulation of  
RT tissue expression by cholesterol feeding."  
RT Biosci. Biotechnol. Biochem. 61:286-290(1997).  
RL

AN  
 RP SEQUENCE OF 25-60.  
 RX MEDLINE=93213845; PubMed=8461329;  
 RA Oku H., Ishikawa M., Nagata J., Toda T., Chinen I.;  
 RT "Lipoprotein and apoprotein profile of Japanese quail.";  
 RL Biochim. Biophys. Acta 1167:22-28(1993).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECTITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF VLDL, HDL, LDL AND IN  
 CC CHYLOMICRONS. EXPRESSED IN A NUMBER OF TISSUES INCLUDING LIVER,  
 CC SMALL INTESTINE, LUNG, KIDNEY, HEART AND MUSCLE WITH HIGHEST  
 CC EXPRESSION IN LIVER AND SMALL INTESTINE.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL: D85133; BAA12729.1; -.  
DR HSSP: P02647; 1GW4.  
DR InterPro: IPR000074; Apolipoprotein.  
KW Pfam: PF01442; Apolipoprotein; 1.  
Plasma: lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.

FT	CHAIN	25	264		APOLIPOPROTEIN A-I.
FT	DOMAIN	67	264		10 X APPROXIMATE TANDEM REPEATS
FT	REPEAT	87	88		1.
FT	REPEAT	89	110		2.
FT	REPEAT	111	121		3 (HALF-LENGTH).
FT	REPEAT	122	143		4.
FT	REPEAT	144	165		5.
FT	REPEAT	166	187		6.
FT	REPEAT	188	209		7.
FT	REPEAT	210	231		8.
FT	REPEAT	232	242		9 (HALF-LENGTH).
FT	REPEAT	243	264		10.
SQ	SEQUENCE	264 AA;	30753 MW;		8781DE213C3F863F CRC64;

Query Match	52.0%;	Score 328;	DB 1;	Length 264;
Best Local Similarity	50.8%;	Pred. No. 1.1e-19;		
Matches 60;	Conservative 29;	Mismatches 29;	Indels 0;	Gaps 0;

[illegible]

RESULT	12		
APAL_ANAPL			
ID	APAL_ANAPL	STANDARD;	PRT; 264 AA.
AC	042296;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	Apolipoprotein A-I precursor (Apo-AI).		
GN	APoA1.		
OS	Anas platyrhynchos (Domestic duck).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.		
OX	NCBI_TaxID=8839;		

```

RP SEQUENCE FROM N.A.
RC STRAIN=PEKIN BREED; TISSUE=Liver;
RA Chen B., Lu X., Wang K., Xue H., Zeng W., Zhu D., Chi L.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -! SUBCELLULAR LOCATION: Extracellular.
CC -! TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -! SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```



RT "Structure, evolution, and polymorphisms of the human apolipoprotein  
RT A4 gene (APOA4).";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86296629; PubMed=3755616;  
RA Karathanasis S.K., Yunis I.;  
RT "Structure, evolution, and tissue-specific synthesis of human  
RT apolipoprotein AIV.";  
RL Biochemistry 25:3962-3970(1986).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87250378; PubMed=3036793;  
RA Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,  
RT Gordon J.I., Taylor J.M.;  
RT "Structure and expression of the human apolipoprotein A-IV gene.";  
RL J. Biol. Chem. 262:7973-7981(1987).  
RN [5]  
RP SEQUENCE OF 21-396 FROM N.A.  
RX MEDLINE=86111885; PubMed=3080432;  
RA Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;  
RT "The nucleotide and derived amino acid sequence of human  
RT apolipoprotein A-IV mRNA and the close linkage of its gene to the  
RT genes of apolipoproteins A-I and C-III.";  
RL J. Biol. Chem. 261:1998-2002(1986).  
RN [6]  
RP SIGNAL SEQUENCE CLEAVAGE SITE.  
RX MEDLINE=84161950; PubMed=6706947;  
RA Gordon J.I., Bisgaler C.L., Sims H.F., Sachdev O.P., Glickman R.M.,  
RT Strauss A.W.;  
RT "Biosynthesis of human preapolipoprotein A-IV.";  
RL J. Biol. Chem. 259:468-474(1984).  
RN [7]  
RP REVIEW ON POLYMORPHISM.  
RA Lohse P., Brewer H.B. Jr.;  
RT "Genetic polymorphism of apolipoprotein A-IV.";  
RL Curr. Opin. Lipidol. 2:90-95(1991).  
RN [8]  
RP VARIANT A-IV\*2.  
RX MEDLINE=90277616; PubMed=2351649;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to  
RT nucleotide substitutions in the apolipoprotein A-IV gene.";  
RL J. Biol. Chem. 265:10061-10064(1990).  
RN [9]  
RP VARIANTS A-IV\*0 AND A-IV\*3.  
RX MEDLINE=90324273; PubMed=1973689;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for  
RT two rare variants of apolipoprotein A-IV-1.";  
RL J. Biol. Chem. 265:12734-12739(1990).  
RN [10]  
RP VARIANTS.  
RX MEDLINE=91310615; PubMed=1677358;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
RT 1(Thr-347-->Ser), apoA-IV-0(Lys-167-->Glu,Gln-360-->His), and apoA-IV-  
RT 3(Glu-165-->Lys).";  
RL J. Biol. Chem. 266:13513-13518(1991).  
RN [11]  
RP ERRATUM.  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RL J. Biol. Chem. 266:19866-19866(1991).  
RN [12]  
RP VARIANT MET-13.  
RX MEDLINE=92238494; PubMed=1349197;  
RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
RT Assmann G.;  
RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
RT gene are associated with changes in the concentration of apo B- and  
RT apo A-I-containing lipoproteins in a normal population.";  
RL Am. J. Hum. Genet. 50:1115-1128(1992).  
RN [13]

RP VARIANT SER-147.  
RX MEDLINE=92144647; PubMed=1737067;  
RA Tenkunen H., Koskinen P., Metsä J., Baumann M., Lukka M.,  
RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
RA Manninen V., Ehnholm C.;  
RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
RT asparagine to serine substitution at residue 127.";  
RL Biochim. Biophys. Acta 1138:27-33(1992).  
RN [14]  
RP VARIANT A-IV\*5.  
RX MEDLINE=93138374; PubMed=1487136;  
RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
RA Ferrell R.E., Pollitzer W.S.;  
RT "Molecular basis of a unique African variant (A-IV 5) of human  
RT apolipoprotein A-IV and its significance in lipid metabolism.";  
RL Genet. Epidemiol. 9:379-388(1992).  
RN [15]  
RP VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
RX MEDLINE=95245341; PubMed=7728150;  
RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
RA Csaszar A.;  
RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
RT frequencies, effect on lipid levels, and sequence of two new  
RT variants.";  
RL Hum. Mutat. 5:58-65(1995).  
RN [16]  
RP VARIANTS FCBL SEATTLE SER-161; LEU-178 AND GIN-264.  
RX MEDLINE=97114287; PubMed=8956036;  
RA Deeb S.S., Nevlin D.N., Iwasaki L., Brunzell J.D.;  
RT "Two novel apolipoprotein A-IV variants in individuals with familial  
RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
RT activity.";  
RL Hum. Mutat. 8:319-325(1996).  
RN [17]  
RP FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
CC COMPONENT OF HDL AND CHYLOMICRONS.  
CC [18]  
RP SUBCELLULAR LOCATION: Extracellular.  
CC [19]  
RP TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
CC [20]  
RP DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC [21]  
RP POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV\*0  
CC TO APOA-IV\*7). APOA-IV\*1 IS THE MAJOR ALLELE (90%), IV\*2 IS ALSO  
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.  
CC [22]  
RP DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN  
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF  
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).  
CC [23]  
RP SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC [24]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC [25]  
RP EMBL: M14642; AAA51745.1; -;  
DR EMBL: X13629; CAA31955.1; -;  
DR EMBL: M14566; AAA51748.1; -;  
DR EMBL: J02758; AAA96731.1; -;  
DR EMBL: M13654; AAA51744.1; -;  
DR PIR: A26481; LPHU04.  
DR PIR: A24449; A24449.  
DR PIR: A29330; A29330.  
DR PIR: A26280; A26280.  
DR PIR: S02715; S02715.  
DR HSP: P02649; INFO.





Sun Sep 22 12:10:25 2002

us-09-803-918a-2\_copy\_25\_144.rsp

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:21:36 ; Search time 111.17 Seconds  
(without alignments)  
186.736 Million cell updates/sec

Title:	US-09-803-918A-2_COPY_25_144
Perfect score:	631
Sequence:	1 DEPPQSPWDRVKDLATVYD.....LDDEQKKWQEEMLYRQKVE 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```

Searched:      562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

```
Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriapi:*
17: sp_archaeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query			ID		Description
No.	Score	Match	Length	DB	ID	
1	508	80.5	264	11	Q9Z2L4	Q9Z2L4 mesocricetu
2	481	76.2	263	11	Q09042	Q09042 mus musculu
3	481	76.2	263	11	Q08855	Q08855 mus musculu
4	464	73.5	258	11	Q09054	Q09054 rattus norv
5	464	73.5	258	11	Q08877	Q08877 rattus norv
6	411	65.1	241	6	Q9TS49	Q9TS49 erinaceus e
7	247	39.1	56	6	Q02762	Q02762 ovis aries
8	217	34.4	82	6	Q29248	Q29248 sus scrofa
9	141	22.3	26	4	Q9UCT8	Q9UCT8 homo sapien
10	139	22.0	174	13	Q9DFQ3	Q9DFQ3 gillichthys
11	125.5	19.9	366	13	Q93601	Q93601 gallus gall
12	119.5	18.9	244	4	Q13784	Q13784 homo sapien
13	114	18.1	263	13	Q98TG6	Q98TG6 anguilla ja
14	114	18.1	435	11	Q01488	Q01488 mus musculu
15	110	17.4	395	11	Q91XF8	Q91XF8 mus musculu
16	108.5	17.2	259	13	Q98TG5	Q98TG5 anguilla ja

17	107	17.0	395	11	Q9DBN0	Q9dbn0 mus musculus
18	106	16.8	34	13	Q9PAR6	Q9prt6 anser anser
19	105.5	16.7	259	13	Q98TG4	Q98tg4 anguilla ja
20	104.5	16.6	259	13	Q98TG1	Q98tg1 anguilla ja
21	103.5	16.4	174	13	Q98S13	Q98s13 cyprinus ca
22	102.5	16.2	259	13	Q98TG3	Q98tg3 anguilla ja
23	101	16.0	30	11	Q9QV04	Q9qv04 rattus sp.
24	100.5	15.9	259	13	Q98TG2	Q98tg2 anguilla ja
25	100.5	15.9	363	4	Q9UBJ3	Q9ubj3 homo sapien
26	99	15.7	313	6	Q9GLC0	Q9glc0 tupala glis
27	96.5	15.3	275	13	Q9PT02	Q9pt02 oncorhynchus
28	91	14.4	178	2	P72564	P72564 streptococc
29	90	14.3	138	2	Q9S613	Q9s613 streptococc
30	89	14.1	1138	5	Q22276	Q22276 caenorhabdi
31	88.5	14.0	254	2	Q9KHL8	Q9khl8 enterococcu
32	88.5	14.0	1814	5	Q9BLM9	Q9blm9 toxocara ca
33	88	13.9	317	6	Q9GLM8	Q9glm8 gorilla gor
34	87.5	13.9	1174	4	Q96D05	Q96dq5 homo sapien
35	87	13.8	317	6	Q9GJU3	Q9gju3 pan troglod
36	87	13.8	317	6	Q9GLM7	Q9glm7 pongo pygma
37	87	13.8	317	6	Q9GLM6	Q9glm6 hylobates l
38	86.5	13.7	911	3	Q06704	Q06704 saccharomyc
39	86.5	13.7	1046	11	Q9Z0H8	Q9z0h8 mus musculu
40	86.5	13.7	1047	11	Q9EP81	Q9ep81 mus musculu
41	86	13.6	336	10	Q9T0F4	Q9t0f4 arabidopsist
42	86	13.6	583	10	Q9ZQC5	Q9zqc5 arabidopsist
43	85	13.5	662	5	Q61096	Q61096 trypanosoma
44	85	13.5	967	5	Q9NGX2	Q9ngx2 entamoeba h
45	85	13.5	980	5	Q96246	Q96246 plasmodium

## ALIGNMENTS

RESULT	1
ID	Q922L4 PRELIMINARY; PRT; 264 AA.
AC	Q922L4; Q922L4;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	APOLIPOPROTEIN A-I.
GN	APOAI.
OS	Mesocricetus auratus (Golden hamster).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC	Mesocricetus.
OX	NCBI_TaxID=10036;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;
RX	MEDLINE=99061559; PubMed=9843713;
RA	Wu J.Y.J., Reeves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;
RT	"Zinc deficiency decreases plasma level and hepatic mRNA abundance of
RT	apolipoprotein A-I in rats and hamsters."
RL	Am. J. Physiol. 275:C1516-C1525(1998).
DR	EMBL: AF046919; AAC98484.1; -.
DR	HSSP: P02647; LAVI.
DR	InterPro: IPR000074; Apolipoprotein.
DR	Pfam: PF01442; Apolipoprotein; 1.
KW	lipoprotein.
SQ	SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;

Query Match	80.5%;	Score 508;	DB 11;	Length 264;
Best Local Similarity	79.0%;	Pred. No. 5.6e-36;		
Matches 94;	Conservative 13;	Mismatches 12;	Indels 0;	Gaps 0.

```
QY      2  EPPQSPMDRVRKDLATVYVVDVKLDSGRDYVSQFEESALGKQLNLKLLDNWDSVTSTFEKRL 61
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      25  DDQPTPMDRVRKDFATVYVDKAVKDSGREYVSQFEESALGKQLNLNLEWMDTLGSTVGRLO 84
QY      62  EQLGPVIOEFWNDLNEKETEGLRQEMSKDLLEVRKAKVQPYLDDFOKKWQOEMELRYRQKVE 120
```

```

      85  FÖLGPVIOEFDNLEKETEMLRKMNKDLEEVKAKVOPYLDQFÖTQWOEEVALYRÖKME 143
      DB

```

RESULT	2		
ID	009042		
AC	009042:	PRELIMINARY;	PRT; 263 AA.
DT	01-JUL-1997	(TREMBLrel. 04, Created)	
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)	
DE	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
OS	APOLIPROTEIN A-I.		
OC	Mus musculus (Mouse).		
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
RN	NCBI_Taxid=10090;		
RC	[1]		
RP	SEQUENCE FROM N.A.		
RA	STRAIN=ICR AND BALB/C, AND C3H; TISSUE=SPLEEN;		
RL	Chiang A.N., Fan R.C., Shaw G.C., Yang U.C.;		
DR	Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.		
DR	EMBL: U79575; AAB58427.1; -		
DR	EMBL: U79572; AAB58424.1; -		
DR	EMBL: U79573; AAB58425.1; -		
DR	HSSP: P02647; IAV1.		
DR	InterPro: IPR000074; Apolipoprotein.		
DR	Pfam: PF01442; Apolipoprotein; 1.		
KW	Lipoprotein.		
SO	SEQUENCE	263 AA; 30516 MW; E245DF7483A5A0DD CRC64;	

Query Match	76.28;	Score 481;	DB 11;	Length 263;
Best Local Similarity	72.9%;	Pred. No. 1.2e-33;		
Matches	86;	Conservative 21;	Mismatches 11;	

```
QY 2 EPPQSPWDRVMDLATVYVDVLKDSGRDVSQFEGSALGQUNLKLDDNWDSTVSTFSKLR 61
   : ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 DEPOQOWMDKXKVFANVYVDKVKDSGRDVSQFESSLGGQUNLLENWDITLGSVSOLO 84

QY 62 EQLGPVTOEFWDNLKETEGTLROENSKDLEEVKAKVQPYLDDFQKKWOEEMETLYROKV 119
   | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 85 ERLGPITTRFEWDLNEKETDWMVRQENKNKDLLEEVKQKVQPYLDEFQKKWKEDEVLYLROKV 142
```

RESULT	3		
ID	008855	PRELIMINARY;	PRT; 263 AA.
AC	008855;		
DT	01-JUL-1997	(TREMBLrel. 04, Created)	
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	APOLIPOPROTEIN A-I.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6; TISSUE=SPLEEN;		
RA	MEDLINE=98077648; PubMed=9415807;		
RT	Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;		
RT	"Repetitive elements in the third intron of murine apolipoprotein A-I		
RT	gene.";		
RL	Biochem. Mol. Biol. Int. 43:989-996(1997).		
DR	EMBL: U79574; AAB58426.1; -.		
DR	HSSP; P02647; IAVI.		
DR	InterPro; IPR000074; Apolipoprotein.		
DR	Pfam; PF01442; Apolipoprotein; 1.		
KW	lipoprotein.		
SO	SEQUENCE	263 AA; 30544 MW; 7825DF7483A5B24A CRC64;	

Query Match *	76.28;	Score 481;	DB 11;	Length 263;
---------------	--------	------------	--------	-------------

Best Local Similarity 72.9%; Pred No. 1.2e-33;  
Matches 86; Conservative 21; Mismatches 11; Indels 0; Gaps 0.

```
QY      2  EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGQLNLKLDNWDVSTSFESKLR 61
      :  ||| ||| | ||| : ||| ||| ||| | : ||| ||| ||| ||| : ||| ||| :
Db      25  DEPQSQWDRVKDFANVYVDAYKDSGRDYSQFESSLGQQLNLNLLENMDTLGSTVSQLO 84
```

QY	62	EQLSPVTOEFWNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEEMELRYKV	119
Db	85	ERIGPLTRDFWNLEKETDWVRQEMNNDLEEVKQKVQPYLDEFQKKMKEDVELRYKV	142

RESULT	4	
009054		
ID	009054	PRELIMINARY; PRT; 258 AA.
AC	009054;	
DT	01-JUL-1997	(TREMBLrel. 04, Created)
DT	01-JUL-1997	(TREMBLrel. 04, last sequence update)
DE	01-JUN-2001	(TREMBLrel. 17, last annotation update)
OS	APOLIPOPROTEIN A-I.	
OC	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=WKY, AND SHRSP; TISSUE=SPLEEN;	
RA	Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;	
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: U79578; AAB58430.1; -.	
DR	EMBL: U79577; AAB58429.1; -.	
DR	HSSP: P02647; IAV1.	
DR	InterPro: IPR000074; Apolipoprotein.	
DR	Pfam: PF01442; Apolipoprotein; 1.	
KW	Lipoprotein.	
SQ	SEQUENCE 258 AA; 29918 MW; 093E6EF2E629C0C8 CRC64;	

Query Match	73.5%;	Score 464;	DB 11;	Length 258;
Best Local Similarity	72.5%;	Pred. No. 3.2e-32;		
Matches 87;	Conservative 14;	Mismatches 17;	Indels 3;	Canc 1.

```
QY 1 DEPPQSPMDRVKDLATVYVYDVCLKDSGRDVSQFEGSALGQNLKLLDNWDSVTSTFFSKL 60
    ||| ||||| ||||| : ||||| ||||| ||||| ||||| : ||| :|
Db 25 DEPSQ--WDRVKDFATVYVDVAVKDSGRDVSQFESSTIGKQLNLNLLDNWDTLGSITVGR 82
```

QY 61 REQLGPVTOEWMNDLEKETEGTLRQESKDL EEVAKAVQPYLDDFOKKWQEEEMELYRQKE 120  
 : ||||| : ||||| : ||||| | : ||||| : ||||| : ||||| :  
 Db 83 QEOLGVPVTOEFWANLEKETDWP RNENMKDL ENVKQKWPHIDFEQEKWNNEVEAYRKOLE 142

RESULT	5	
ID	008877	PRELIMINARY; PRT; 258 AA.
AC	008877;	
DT	01-JUL-1997	(TREMBLrel. 04, Created)
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DE	APOLIPOPROTEIN A-I.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SHR; TISSUE=SPLEEN;	
RX	MEDLINE=98077648; PubMed=9415807;	
RA	Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;	
RT	"Repetitive elements in the third intron of murine apolipoprotein A-I	
RT	gene.";	
RL	Biochem. Mol. Biol. Int. 43:989-996(1997).	
DR	EMBL: U79576; AAB58428.1; -.	
DR	HSSP: P02647; IAVI.	

DR PROCEM; MOL. BIOL. INT. 43:989-996(1997).  
DR EMBL; U79576; AAB58428.1; -  
DR HSSP; P02647; LAV1.





RT "Identification of apolipoprotein A1 and immunoglobulin as components  
of a serum complex that mediates activation of human sperm motility.";  
RL Biochemistry 30:8986-8990(1991).  
SQ SEQUENCE 26 AA; 2930 MW; 64921A333E768D27 CRC64;

Query Match 22.3%; Score 141; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 9.6e-06;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSG 26  
DB 1 DEPPQSPWDRVKDLATVYVDVLKDSG 26

RESULT 10

Q9DFQ3 PRELIMINARY; PRT; 174 AA.  
AC Q9DFQ3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOLIPROTEIN A-IV (FRAGMENT).  
OS Gillichthys mirabilis (Long-jawed mudsucker).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidae;  
OC Gobiidae; Gillichthys.  
OX NCBI\_TaxID=8222;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=21117151; PubMed-11172064;  
RA Gracey A.Y., Troll J.V., Somero G.N.;  
RT "Hypoxia-induced gene expression profiling in the euryoxic fish  
Gillichthys mirabilis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).  
DR EMBL; AF266179; AAG13299.1;  
KW Lipoprotein.  
FT NON\_TER 174 174  
SQ SEQUENCE 174 AA; 19342 MW; 3440836F5EDC863B CRC64;

Query Match 22.0%; Score 139; DB 13; Length 174;  
Best Local Similarity 27.18%; Pred. No. 0.00015;  
Matches 32; Conservative 25; Mismatches 61; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNKLNDWDSVTSTFSKLR 61  
DB 21 EQPKTBLDLVRDFTFWDYVAKATQMAEGSKQIQDSSELGQEVNTKLSQSANTYVXQYVALK 80  
QY 62 EQLGPTVTOEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKV 119  
DB 81 TQAGPLSQDLTLQITQCTETLKSRLLETLSAVNANLQPYAEQLVADLTQTKMDALKXDY 138

RESULT 11  
Q93601 PRELIMINARY; PRT; 366 AA.

AC Q93601;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE APOLIPROTEIN AIV.  
GN APOAIV.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98221191; PubMed=9553114;

RA Steinmetz A., Hermann M., Nimf J., Aebersold R., Ducret A.,  
RA Weinberg R.B., Schneider W.J.;  
RT "Expression and conservation of apolipoprotein AIV in an avian  
RT species.";  
RL J. Biol. Chem. 273:10543-10549(1998).  
DR EMBL; Y16534; CAA76273.1;  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 366 AA; 40852 MW; B47DB49F3E8D91CE CRC64;

Query Match 19.9%; Score 125.5; DB 13; Length 366;  
Best Local Similarity 29.4%; Pred. No. 0.0053;  
Matches 30; Conservative 26; Mismatches 45; Indels 1; Gaps 1;

QY 6 SPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNKLNDWDSVTSTFSKLRQDLG 65  
DB 22 SP-DQVATVLMWRYFTELGSSAKETVDDLOQAEITKQLNTLLQSNLQSANSAEELQRLV 80

QY 66 PVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKK 107  
DB 81 PFATELQALVQDSQRLKQIQQLAEIQAALPYADEVHQ 122

RESULT 12

Q13784 PRELIMINARY; PRT; 244 AA.  
AC Q13784;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOA4 PROTEIN (FRAGMENT).  
GN APOA4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=86016704; PubMed=3931073;  
RA Karathanasis S.K.;  
RT "Apolipoprotein multigene family: tandem organization of human  
RT apolipoprotein A1, CIII, and AIV genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:6374-6378(1985).  
DR EMBL; M10373; AAB59516.1;  
DR HSSP; P02649; 1B24.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 244 AA; 28157 MW; 01616ACB53F12734 CRC64;

Query Match 18.9%; Score 119.5; DB 4; Length 244;  
Best Local Similarity 25.98%; Pred. No. 0.011;  
Matches 30; Conservative 31; Mismatches 46; Indels 9; Gaps 2;

QY 11 VKDLATVYVDVLKDSGRDYSQFEGSALGKQLNKLNDWDSVTSTFSKLRQDLGPTQE 70  
DB 82 VEELRRSLAPYAQDTQEKLNHOLEGLTFQMKNAEELK--ARISASAEELRORLAPLAE 139  
QY 71 FWDNLEKETEGLRQEMS-----KDLFEVAKAVQPYLDDFQKKWQEMELYRQKV 119  
DB 140 VRGNLKGNTGSLAELGHLDDQVEEFRRRVEPYGENFNKALVQOMQDLRQKL 195

RESULT 13

Q98TG6 PRELIMINARY; PRT; 263 AA.  
AC Q98TG6;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

```

QY      1  DEP-POSPPWDRYKDLATVYVVDVLKDSGRDVSQFEGSALGOKLNLK---LDN----- 49
      | | | | : : : : : : : : : : : : : : | | | |
Db      24  DAPAPPSOLEHVRAAVGMYLQOVKEFAQALEHLDDEY- KDYKRLRSQSLDNIGGYIOS 82
QY      50  -----WDSVTSTFSKLRQGLGPTVQFEWDNLKEKETGLROEM----SKDLEEVKAV 97
      | : | : | : : : : : | | : | : : : : | : | : | :
Db      83  ASAALSPYTDVAVSSQFMELTKDMRDKIQADVDQLKDLQPKDELKEVYQKHLEDEYRAKL 142
QY      98  QPYLDDFOKKWQEMELLYRQKVE 120
      : | : : : : : | : : : : : : : : : : : :
Db     143  EPLVKEYTEKNHQEMELKTKLQ 165

```

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE VARIANT APOLIPROTEIN A-IV  
NN APOA-4  
APPO-4

J. Biol. Chem. 265:12715-12721(1991).

-1- TISSUE SPECIFICITY: INTESTINE, LIVER, AND PLASMA.

MEMBRANES: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A & B, RELATED 11-MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-HELICAL, & MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES. MISCELLANEOUS: THE APOA-IV POLYMORPHISMS ARE CLASSIFIED BASED ON

Query Match	18.18;	Score 114;	DB 11;	Length 435;
Best Local Similarity	21.48;	Pred. No. 0.064;		
Matches 24;	Conservative 37;	Mismatches 51;	Indels 0;	Cars 0

RESULT	15	
Q91XF8		
ID	Q91XF8	PRELIMINARY;
CC	Q91XF8	PRT;
		395 AA

NCBI\_TaxID=10090; Mus  
Sciurognathi; Muridae; Mus  
Rodentia; Sciurognathi; Muridae; Mus  
Mammalia; Eutheria; Rodentia; Muridae; Mus  
Mammalia; Eutheria; Rodentia; Muridae; Mus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Muridae; Mus

Query Match	17.4%;	Score 110;	DB 11;	Length 395;
Best Local Similarity	20.5%;	Pred. No. 0.12;		
Matches 23;	Conservative 38;	Mismatches 51;	Indels 0;	Cons 0

QY 9 DRVKDLATVYVDVLKDSGRDYVSQEEGSA LGKQLNLKLLDNWSDVTSTFSKLREQLPVT 68  
| : | :: | : : : : | | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :  
Db 25 DQVANVWDYFTQLSNNAKAEVQFQKTDTVQTQLSTLFQDKLGDASTYADGVHNKKLVPEV 84  
  
QY 69 QEFWNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQVE 120  
: : | : | : : : | : | : : : : | : : : : : : | : : : : : : | : : : : : :  
Db 85 VQLSGHLAQETERVKKEIKLEDLD RMMPHANKYTQTGFENMQKLQEHK 136

Search completed: September 22, 2002, 12:21:37  
Job time: 1117 sec

Sun Sep 22 12:10:26 2002

us-09-803-918a-2\_copy\_25\_144.rspt

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:06:28 ; Search time 66.91 Seconds  
(without alignments)  
244.137 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_25\_194  
Perfect score: 885  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....RARAHVDALRTHLAPYSDEL 170

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	885	100.0	267	1	LPHUA1	apolipoprotein A-I
2	842	95.1	267	1	A26529	apolipoprotein A-I
3	842	95.1	267	2	JS0079	apolipoprotein A-I
4	758	85.6	266	1	LPDGA1	apolipoprotein A-I
5	737	83.3	265	2	A46018	apolipoprotein AI
6	730	82.5	265	2	UT0672	apolipoprotein A-I
7	728.5	82.3	264	2	S31394	apolipoprotein A-I
8	711	80.3	266	1	LPRB12	apolipoprotein A-I
9	708	80.0	265	2	A56858	apolipoprotein A-I
10	704	79.5	265	1	LPRB1B	apolipoprotein A-I
11	701.5	76.7	241	2	A24998	apolipoprotein A-I
12	678.5	76.7	231	2	J00704	apolipoprotein A-I
13	655	74.0	262	2	JC1237	apolipoprotein A-I
14	655	74.0	264	2	S22420	apolipoprotein A-I
15	573	64.7	259	2	A24700	apolipoprotein A-I
16	471	53.2	246	2	A61448	apolipoprotein A-I
17	471	53.2	264	1	LPEHA1	apolipoprotein A-I
18	454	51.3	264	2	JC5456	apolipoprotein A-I
19	370	41.8	164	2	S21830	apolipoprotein A-I
20	214	24.2	429	2	S29565	apolipoprotein A-I
21	213	24.1	396	1	LPHUA4	apolipoprotein A-I
22	210	23.7	401	2	A47141	apolipoprotein A-I
23	181	20.5	391	1	LPRUA4	apolipoprotein A-I
24	175	19.8	391	2	B40892	apolipoprotein A-I
25	175	19.8	394	2	A25281	apolipoprotein A-I
26	175	19.8	395	2	A40892	apolipoprotein A-I
27	175	19.8	399	2	C40892	apolipoprotein A-I
28	156.5	17.7	311	2	JU0036	apolipoprotein E p
29	156.5	17.7	312	1	LPRTE	apolipoprotein E p

30	156	17.6	258	2	JH0472	apolipoprotein A-I
31	141.5	16.0	329	2	JC5566	apolipoprotein E p
32	138.5	15.6	317	2	S33450	apolipoprotein E -
33	135.5	15.3	291	2	C60940	apolipoprotein E -
34	133	15.0	298	2	S12635	apolipoprotein E p
35	125	14.1	317	2	A28792	apolipoprotein E p
36	125	14.1	317	2	S03185	apolipoprotein E p
37	119.5	13.5	317	1	LPHUE	apolipoprotein E p
38	119.5	13.5	896	2	S43074	epidermal growth f
39	119.5	13.5	1547	2	T28657	blackjack protein,
40	118.5	13.4	316	2	JC6549	apolipoprotein E p
41	118.5	13.4	897	2	A54696	EGF receptor subst
42	115	13.0	180	2	C70428	hypothetical prote
43	113	12.8	316	2	S26478	apolipoprotein E -
44	113	12.8	316	2	I45996	apolipoprotein E -
45	113	12.8	470	2	T23512	hypothetical prote

## ALIGNMENTS

RESULT 1  
LPHUA1  
apolipoprotein A-I precursor [validated] - human  
N/Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C/Species: Homo sapiens (man)  
C/Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C/Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010; 6197  
R;Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A/Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A/Reference number: A90947; MUID:85026665  
A/Accession: A90947  
A/Molecule type: DNA  
A/Residues: 1-267 <SEI>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
A/Accession: B90947  
A/Molecule type: mRNA  
A/Residues: 1-267 <SEI>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R;Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988  
A/Title: Sequence and expression of tangler apoA-I gene.  
A/Reference number: S02373; MUID:88196137  
A/Accession: S02373  
A/Molecule type: DNA  
A/Residues: 1-267 <MAK>  
A/Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
R;Shoulders, C.C.; Kornblitt, A.R.; Munro, B.S.; Baralle, F.E.  
Nucleic Acids Res. 11, 2827-2837, 1983  
A/Title: Gene structure of human apolipoprotein AI.  
A/Reference number: A93465; MUID:83220822  
A/Accession: A93465  
A/Molecule type: DNA  
A/Residues: 1-267 <SHO>  
A/Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518;  
R;Kathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
A/Title: Isolation and characterization of the human apolipoprotein A-I gene.  
A/Reference number: A21147; MUID:84016011  
A/Accession: A21147  
A/Molecule type: DNA  
A/Residues: 1-267 <KAR>  
A/Cross-references: GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.  
Nucleic Acids Res. 12, 3917-3932, 1984  
A/Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundant  
A/Reference number: A93519; MUID:84221405  
A/Accession: A93519  
A/Molecule type: mRNA  
A/Residues: 1-267 <SHA>  
A/Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753



A:Accession: B93519  
A:Molecule type: DNA  
A:Residues: 1-24 <SH2>  
R:Cheung, P.; Chan, L.  
Nucleic Acids Res. 11, 3703-3715, 1983  
A:Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
A:Reference number: A93472; MUID:83220772  
A:Accession: A93472  
A:Molecule type: mRNA  
A:Residues: 1-267 <CHE>  
A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GEE  
R:Law, S.W.; Brewer Jr., H.B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
A:Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
A:Reference number: A94010; MUID:84119464  
A:Accession: A94010  
A:Molecule type: mRNA  
A:Residues: 1-267 <LAM>  
A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GEE  
R:Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
A:Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
A:Reference number: A21118; MUID:83195100  
A:Accession: A21118  
A:Molecule type: mRNA  
A:Residues: 1-24 <ZAN>  
R:Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
A:Title: Human plasma proapoA-I: isolation and amino-terminal sequence.  
A:Reference number: A90112; MUID:83256553  
A:Accession: A90112  
A:Molecule type: protein  
A:Residues: 19-27 <BRE>  
R:Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
A:Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high c  
A:Reference number: A90209; MUID:78123731  
A:Accession: A90209  
A:Molecule type: protein  
A:Residues: 25-57, 'Q', 59-169, 'OO', 172-267 <BR2>  
R:Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
J. Clin. Invest. 82, 803-807, 1988  
A:Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A  
A:Reference number: A30516; MUID:88331387  
A:Accession: A30516  
A:Molecule type: protein  
A:Residues: 25-56 <YUI>  
R:Nichols, W.C.; Dwulet, F.E.; Liepnieks, J.; Benson, M.D.  
Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
A:Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid.  
A:Reference number: A31582; MUID:89050104  
A:Accession: A31582  
A:Molecule type: protein  
A:Residues: 25-49, 'R', 51-85, 'D', 87-107 <NIC>  
A:Note: variant sequence from patient with familial amyloidotic polyneuropathy type III  
R:Manjunath, P.; Marcel, Y.L.; Uma, J.; Seldah, N.G.; Chretien, M.; Chapdelaine, A.  
J. Biol. Chem. 264, 16853-16857, 1989  
A:Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
A:Reference number: A34409; MUID:89380318  
A:Accession: A34409  
A:Molecule type: protein  
A:Residues: 25-48 <MAN>  
R:Stoffel, W.; Binczek, E.  
Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
A:Title: Structural requirements of human preproapolipoprotein AI for translocation and  
A:Reference number: S02737; MUID:89149957  
A:Accession: S02737  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-6, 'AV', 9, 'LV', 12-29 <STO>  
A:Note: part of this sequence, including the amino end of the mature protein, was confir  
R:Stoffel, W.; Binczek, E.  
Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A;Title: Transient expression of wild type and mutant human apolipoprotein AI in COS  
A;Reference number: S16197; MUID:92029676  
A;Contents: annotation;extension of studies in reference S02737  
R;Stoffel, W.; Kruger, E.; Deutzmann, R.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
A;Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processed  
A;Reference number: A19913; MUID:83236195  
A;Accession: B19913  
A;Molecule type: protein  
A;Residues: 1-6,'X',8-13,'XXX',17-18,'XX',21,'X',23-25,'X',27-29 <ST2>  
R;Einholm, C.; Bozas, S.E.; Tenkanen, H.; Kitzsbaum, L.; Metsu, J.; Murphy, B.; Walke  
Biochim. Biophys. Acta 1086, 255-260, 1991  
A;Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
A;Reference number: A56815; MUID:92075698  
A;Accession: A56815  
A;Molecule type: protein  
A;Residues: 25-31,'P',33 <EHN>  
A;Experimental source: serum  
A;Note: sequence extracted from NCBI backbone (NCBIP:69759)  
A;Note: 32-Tip was also found  
R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
Biochemistry 33, 1988-1993, 1994  
A;Title: Identification of proteins associated with apolipoprotein A-I-containing lip  
A;Reference number: A54223; MUID:94162201  
A;Accession: A54223  
A;Molecule type: protein  
A;Residues: 25-39 <KUN>  
R;Mogulievsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.;  
DNA 8, 429-436, 1989  
A;Title: Production of human recombinant proapolipoprotein A-I in *Escherichia coli*: p  
A;Reference number: I39476; MUID:89377481  
A;Accession: I39476  
A;Molecule type: mRNA  
A;Residues: 19-267 <RES>  
A;Cross-references: GB:M29068; NID:g178774; PIDN:AAA51747.1; PID:g178775  
R;Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B.  
J. Biol. Chem. 263, 18530-18536, 1988  
A;Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by th  
A;Reference number: I39475; MUID:89054040  
A;Accession: I39475  
A;Molecule type: DNA  
A;Residues: 1-14 <RE2>  
A;Cross-references: GB:J04066; NID:g178763; PIDN:AAA51746.1; PID:g553183  
R;Breslow, J.L.  
Ann. Rev. Biochem. 54, 699-727, 1985  
A;Title: Human apolipoprotein molecular biology and genetic variation.  
A;Reference number: A90042; MUID:85278004  
A;Contents: annotation; review of sequences, variants and gene location  
R;Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B.  
J. Biol. Chem. 261, 3911-3914, 1986  
A;Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acyl  
A;Reference number: A92577; MUID:86140194  
A;Contents: annotation; acylation with palmitate  
A;Note: an undetermined serine or threonine is acylated by fatty acid; the acylating  
R;Law, S.W.; Brewer, H.B.  
J. Biol. Chem. 260, 12810-12814, 1985  
A;Title: Tangier disease: The complete mRNA sequence encoding for preapo-A-I.  
A;Reference number: I55236; MUID:86008382  
A;Accession: I55236  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-143,'D',145-267 <RE3>  
A;Cross-references: GB:M11791; NID:g178776; PIDN:AAA35545.1; PID:g178777  
C;Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intesti  
oteins (HDL) in plasma.  
C;Genetics:  
A;Gene: GDB:APOA1  
A;Cross-references: GDB:119684; OMIM:107680  
A;Map position: 11q23.3-11q23.3  
A;Introns: 15/1, 67/2  
C;Function:  
A;Description: participates in the reverse transport of cholesterol from tissues to t  
sterol acyltransferase (LCAT); noncovalently binds and stabilizes prostacyclin (Pgt-2

C:Superfamily: apolipoprotein A-I  
C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipi  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-267/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 100.0%; Score 885; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 4.7e-49;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
|||||  
Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
QY 61 REOLGPVTQEEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVE 120  
|||||  
Db 85 REOLGPVTQEEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVE 144  
QY 121 PLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDEL 170  
|||||  
Db 145 PLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDEL 194

## RESULT 2

apolipoprotein A-I precursor - crab-eating macaque

C:Species: Macaca fascicularis (crab-eating macaque)  
C:Date: 30-Sep-1989 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999

C:Accession: A26529; A26627; S23135; A57766

R:Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marotti, K.R.

A:Title: The primary structure of cynomolgus monkey apolipoprotein A-1 deduced from the

A:Reference number: A26529; MUID:87191989

A:Accession: A26529

A:Molecule type: mRNA

A:Residues: 1-267 <POL>

A:Cross-references: GB:M15411; NID:g342074; PIDN:AAA36834.1; PID:g342075

R:Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolosi,

Biochemistry 26, 1457-1463, 1987

A:Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (cyn

A:Reference number: A26627; MUID:87185451

A:Accession: A26627

A:Molecule type: protein

A:Residues: 25-48 <HER>

R:Murray, R.W.; Marotti, K.R.

Biochim. Biophys. Acta 1131, 207-210, 1992

A:Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and corres

A:Reference number: S23135; MUID:92305062

A:Accession: S23135

A:Molecule type: DNA

A:Residues: 1-12, 'L', 14-267 <MUR>

A:Cross-references: GB:M83242; NID:g342070; PIDN:AAA36832.1; PID:g342071

R:Sorci-Thomas, M.; Kearns, M.W.

J. Biol. Chem. 266, 18045-18050, 1991

A:Title: Transcriptional regulation of the apolipoprotein A-I gene.

A:Reference number: A57766; MUID:92011532

A:Accession: A57766

A:Molecule type: DNA

A:Residues: 1-10 <RES>

A:Cross-references: GB:M69223; NID:g342066; PIDN:AAA36831.1; PID:g553820

C:Comment: The precursor is synthesized in the liver and small intestine. The propeptide

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

Query Match 95.1%; Score 842; DB 1; Length 267;  
Best Local Similarity 94.1%; Pred. No. 2.4e-46;  
Matches 160; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
|||||  
Db 25 DEPPQTPWDRVKDLATVYVEALKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
QY 61 REOLGPVTQEEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVE 120  
|||||  
Db 85 REOLGPVTQEEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVE 144  
QY 121 PLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDEL 170  
|||||  
Db 145 PLRAELHEGTRQKLHELHEKLSPLGEEVRDRARAHVDALRTHLAPYSDEL 194

## RESULT 3

apolipoprotein A-I precursor - baboon

C:Species: Papio sp. (baboon)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 13-Jun-1997

C:Accession: JS0079

R:Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; Vandenberg, J.L.

Gene 74, 483-490, 1988

A:Title: The baboon gene for apolipoprotein A-I: characterization of a cDNA clone and

A:Reference number: JS0079; MUID:89232739

A:Accession: JS0079

A:Molecule type: mRNA

A:Residues: 1-267 <HIX>

A:Experimental source: liver

C:Comment: This protein is the principal protein component of high density lipoprotein

C:Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase

C:Comment: This protein contains a region of repeated amino acids which form amphipat

C:Genetics:

A:Gene: apoA1

C:Superfamily: apolipoprotein A-I

C:Keywords: HDL; lipid binding; lipoprotein

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-267/Product: apolipoprotein A-I #status predicted <LAI>

F:123-144,145-166,167-188,189-210,211-232,233-254/Region: tandem repeats

Query Match 95.1%; Score 842; DB 2; Length 267;  
Best Local Similarity 94.1%; Pred. No. 2.4e-46;  
Matches 160; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
|||||  
Db 25 DEPPQTPWDRVKDLATVYVEALKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
QY 61 REOLGPVTQEEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVE 120  
|||||  
Db 85 REOLGPVTQEEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVE 144  
QY 121 PLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDEL 170  
|||||  
Db 145 PLRAELHEGTRQKLHELHEKLSPLGEEVRDRARAHVDALRTHLAPYSDEL 194

## RESULT 4

apolipoprotein A-I precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 17-Dec-1982 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C:Accession: A60940; A03092; A61418

R:Uno, C.C.; Li, W.H.; Chan, L.

J. Lipid Res. 30, 1735-1746, 1989

A:Title: Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implic

A:Reference number: A60940; MUID:90132271

A:Accession: A60940

A:Molecule type: mRNA

A:Residues: 1-266 <LUO>











Sun Sep 22 12:10:31 2002

us-09-803-918a-2\_copy\_25\_194.rpr

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:22:16 ; Search time 35.02 Seconds  
(without alignments)  
187.959 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_194

Perfect score: 885  
Sequence: 1 DEPPQSPWDRVKDLATYVD.....RARAHVDALRTHLAPYSDEL 170

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	885	100.0	267	1	APAI_HUMAN P02647 homo sapien
2	842	95.1	267	1	APAI_MACFA P15568 macaca fasc
3	758	85.6	266	1	APAI_CANFA P02648 canis fami
4	742	83.8	265	1	APAI_PIG P18648 sus scrofa
5	711	80.3	266	1	APAI_RABIT P09809 oryctolagus
6	703	79.4	265	1	APAI_BOVIN P15497 bos taurus
7	655	74.0	264	1	APAI_MOUSE O00623 mus musculu
8	637	72.0	265	1	APAI_TUPGB O18759 tupai glis
9	573	64.7	259	1	APAI_RAT P04639 rattus norv
10	471	53.2	264	1	APAI_ANAPL O42296 anas platyr
11	471	53.2	264	1	APAI_CHICK P08250 gallus gall
12	454	51.3	264	1	APAI_COTJA P32918 coturnix co
13	214	24.2	429	1	APAI_MACFA P33621 macaca fasc
14	210	23.7	401	1	APAI_PAPAN Q28758 papio anubi
15	209	23.6	396	1	APAI_HUMAN P06727 homo sapien
16	192	21.7	382	1	APAI_PIG O46409 sus scrofa
17	184.5	20.8	262	1	APAI_ONCMY O57523 oncorhynchu
18	182	20.6	281	1	APAI_BRARE O42364 brachydanio
19	181	20.5	391	1	APAI_RAT P02651 rattus norv
20	177.5	20.1	262	1	APAI_SALTR Q91488 salmo trutt
21	175	19.8	262	1	APAI_ONCMY O57524 oncorhynchu
22	175	19.6	395	1	APAI_MOUSE P06728 mus musculu
23	173.5	19.6	262	1	APAI_BRARE O42363 brachydanio
24	156.5	17.7	311	1	APAI_MOUSE P08226 mus musculu
25	156.5	17.7	312	1	APAI_RAT P02650 rattus norv
26	154.5	17.5	260	1	APAI_SPAUV O42175 sparus aur
27	153	17.3	258	1	APAI_SALSA P27007 salmo salar
28	138.5	15.6	317	1	APAI_PIG P18650 sus scrofa
29	135.5	15.3	305	1	APAI_CANFA P18649 canis fami
30	133	15.0	298	1	APAI_CAVPO P23529 cavia porce
31	125	14.1	317	1	APAI_MACFA P10517 macaca fasc
32	125	14.1	317	1	APAI_PAPAN P05770 papio anubi
33	119.5	13.5	317	1	APAI_HUMAN P02649 homo sapien

34	119.5	13.5	896	1	EP15_HUMAN	P42566 homo sapien
35	118.5	13.4	897	1	EP15_MOUSE	P42567 mus musculu
36	115	13.0	180	1	YE76_AQUAE	O67453 aquifex aeo
37	113	12.8	316	1	APAI_BOVIN	O03247 bos taurus
38	111.5	12.6	539	1	MYS3_HYDAT	P39922 hydra atten
39	110.5	12.5	520	1	RECNA_AQUAE	O66834 aquifex aeo
40	109	12.3	311	1	APAI_RABIT	P18287 oryctolagus
41	108	12.2	2230	1	GOG4_HUMAN	Q13439 homo sapien
42	107.5	12.1	1938	1	MYS_AEQIR	P24733 aequipecten
43	106	12.0	429	1	SEP6_MOUSE	Q91t4 mus musculu
44	106	12.0	664	1	LAMA_HUMAN	P02545 homo sapien
45	105.5	11.9	365	1	RECF_CHLTR	O84077 chlamydia t

## ALIGNMENTS

RESULT	ID	STANDARD;	PRT;	267 AA.
APAI_HUMAN	APAI_HUMAN			
AC	P02647;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).			
GN	APOA1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84221405; PubMed=6328445;			
RA	Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,			
RA	Baralle F.E.;			
RT	"Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA			
RT	abundance.";			
RL	Nucleic Acids Res. 12:3917-3932(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8502665; PubMed=6207999;			
RA	Seihamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;			
RA	"Isolation and DNA sequence of full-length cDNA and of the entire			
RT	gene for human apolipoprotein AI -- discovery of a new genetic			
RT	polymorphism in the apo AI gene.";			
RL	DNA 3:309-317(1984).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83220822; PubMed=6406984;			
RA	Shoulders C.C., Kornblith A.R., Munro B.S., Baralle F.E.;			
RT	"Gene structure of human apolipoprotein AI.";			
RL	Nucleic Acids Res. 11:2827-2837(1983).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83220772; PubMed=6304641;			
RA	Cheung P., Chan L.;			
RT	"Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";			
RL	Nucleic Acids Res. 11:3703-3715(1983).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84119464; PubMed=6198645;			
RA	Law S.W., Brewer H.B. Jr.;			
RT	"Nucleotide sequence and the encoded amino acids of human			
RT	apolipoprotein A-I mRNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86008382; PubMed=2995392;			
RA	Law S.W., Brewer H.B. Jr.;			
RT	"Tangier disease. The complete mRNA sequence encoding for			
RT	preproapo-A-I.";			
RL	J. Biol. Chem. 260:12810-12814(1985).			
RN	[7]			



RP SEQUENCE FROM N.A.  
 RX MEDLINE=84016011; PubMed=6413973;  
 RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
 RT "Isolation and characterization of the human apolipoprotein A-I  
 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
 RN [18]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89377481; PubMed=2673706;  
 RA Mogulievsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,  
 RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
 RA Holmquist L., Carlson L.A., Bollen A.;  
 RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
 coli: purification and biochemical characterization.";  
 RL DNA 8:429-436(1989).  
 RN [19]  
 RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
 RX MEDLINE=88196137; PubMed=3129297;  
 RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
 RA Zannis V.I.;  
 RT "Sequence and expression of Tangier apoA-I gene.";  
 RL Eur. J. Biochem. 173:465-471(1988).  
 RN [110]  
 RP SEQUENCE OF 118-267 FROM N.A.  
 RX MEDLINE=83091059; PubMed=6294659;  
 RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
 RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
 RT "Isolation and characterization of cDNA clones for human  
 apolipoprotein A-I.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
 RN [111]  
 RP SEQUENCE OF 19-27.  
 RX MEDLINE=83256553; PubMed=6409108;  
 RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
 RA Light J.A.;  
 RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
 RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
 RN [112]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=78123731; PubMed=204308;  
 RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
 RA Bronzert T.J.;  
 RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated  
 from high density lipoproteins.";  
 RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
 RN [113]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=75133493; PubMed=164450;  
 RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;  
 RT "The primary structure of human plasma high density apolipoprotein  
 glutamine I (APOA-I). II. The amino acid sequence and alignment of  
 cyatogen bromide fragments IV, III, and I.";  
 RL J. Biol. Chem. 250:2725-2738(1975).  
 RN [114]  
 RP SEQUENCE OF 25-56.  
 RX MEDLINE=88331387; PubMed=3047170;  
 RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
 RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
 A-I (Apo A-I). A novel function of Apo A-I.";  
 RL J. Clin. Invest. 82:803-807(1988).  
 RN [115]  
 RP SEQUENCE OF 25-48.  
 RX MEDLINE=89380318; PubMed=2506184;  
 RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
 RA Chapdelaine A.;  
 RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
 proteins.";  
 RL J. Biol. Chem. 264:16853-16857(1989).  
 RN [116]  
 RP SEQUENCE OF 25-43.  
 RX MEDLINE=88070603; PubMed=3120314;  
 RA Price R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
 RA Pereira M.E.A.;  
 RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
 neuraminidase, to high-density lipoprotein.";  
 RL Science 238:1417-1419(1987).  
 RN [117]  
 RP SEQUENCE OF 25-42.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 RN [118]  
 RP PALMITOYLATION.  
 RX MEDLINE=86140194; PubMed=3005308;  
 RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
 RT "Human apolipoprotein A-I. Post-translational modification by fatty  
 acid acylation.";  
 RL J. Biol. Chem. 261:3911-3914(1986).  
 RN [119]  
 RP PROCESSING.  
 RX MEDLINE=83195100; PubMed=6405383;  
 RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
 RA Breslow J.L.;  
 RT "Intracellular and extracellular processing of human apolipoprotein  
 A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
 RN [120]  
 RP STRUCTURE BY NMR OF 190-209.  
 RX MEDLINE=96270776; PubMed=8664326;  
 RA Wang G., Treleaven W.D., Cushley R.J.;  
 RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
 presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
 and CD. Evidence for specific peptide-SDS interactions.";  
 RL Biochim. Biophys. Acta 1301:174-184(1996).  
 RN [121]  
 RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
 RX MEDLINE=98024124; PubMed=9356442;  
 RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
 RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
 lipid-bound conformation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
 RN [122]  
 RP VARIANT MILANO.  
 RX MEDLINE=83109095; PubMed=6401735;  
 RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,  
 RA Franceschini G., Sirtori C.R.;  
 RT "Apolipoprotein A-IMilano. Detection of normal A-I in affected  
 subjects and evidence for a cysteine for arginine substitution in the  
 variant A-I.";  
 RL J. Biol. Chem. 258:2508-2513(1983).  
 RN [123]  
 RP VARIANT TANGIER.  
 RX MEDLINE=83300108; PubMed=6412234;  
 RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;  
 RT "Tangier disease: defective recombination of a specific Tangier  
 apolipoprotein A-I isoform (pro-apo A-I) with high density  
 lipoproteins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).  
 RN [124]  
 RP VARIANT NORWAY.  
 RX MEDLINE=84289383; PubMed=6432779;  
 RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,  
 RA Utermann G., Haas J., Steinmetz A., Menzel H.J., Assmann G.;  
 RT "Abnormal lecithin:cholesterol acyltransferase activation by a human  
 apolipoprotein A-I variant in which a single lysine residue is  
 deleted.";  
 RL J. Biol. Chem. 259:10063-10070(1984).  
 RN [125]  
 RP SEQUENCE OF 25-107 (VARIANT IOWA).  
 RX MEDLINE=89050104; PubMed=3142462;  
 RA Nichols W.C., Dwulet F.E., Liepnieks J., Benson M.D.;  
 RT "Variant apolipoprotein AI as a major constituent of a human  
 hereditary amyloid.";



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=90132271; PubMed=2515239;  
RA Luo C.-C., Li W.-H., Chan L.;  
RT "Structure and expression of dog apolipoprotein A-I, E, and C-II  
RN mRNAs: implications for the evolution and functional constraints of  
RL apolipoprotein structure.";  
RN J. Lipid Res. 30:1735-1746(1989).  
[2]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=82142425; PubMed=6801039;  
RA Chung H., Randolph A., Reardon I., Heinrichson R.L.;  
RT "The covalent structure of apolipoprotein A-I from canine high  
RL density lipoproteins.";  
RN J. Biol. Chem. 257:2961-2967(1982).  
[3]  
RP SEQUENCE OF 25-57 AND 262-265.  
RX MEDLINE=76210910; PubMed=179887;  
RA Nakai T., Whayne T.F., Tang J.;  
RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein  
RL A-I.";  
RN FEBS Lett. 64:409-411(1976).  
[4]  
RP SEQUENCE OF 25-37.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
RL dog heart proteins.";  
RN Electrophoresis 18:2795-2802(1997).  
[5]  
RP -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
RC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
RX CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS, SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
DR PIR: A03092; LPDGA1.  
DR HSSP: P02647; 1ODR.  
DR HSC-2DPAGE; P02648; DOG.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 266  
FT DOMAIN 67 266  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 231  
FT REPEAT 232 242  
FT REPEAT 243 266  
FT CONFLICT 168 168 A -> G (IN REF. 2).  
FT CONFLICT 202 202 E -> Q (IN REF. 2).  
FT CONFLICT 235 235 E -> Q (IN REF. 2).  
FT CONFLICT 264 266 NAQ -> A (IN REF. 3).  
SQ SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;

Query Match 85.6%; Score 758; DB 1; Length 266;  
Best Local Similarity 85.2%; Pred. No. 6.8e-43;  
Matches 144; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDIATVYVDVLKDSGRVVSQFEGSALGKQNLKLLDNWDSVSTFSKLR 61  
Db : |||||  
25 DEQSPWDRVKDIATVYVDVAVKDSGRVVAQFEASALGKQNLKLLDNWDSLSSTVTKLR 84  
QY 62 EQLPVTVQEFMDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQKVEP 121  
Db ||:|||||  
85 EQIGPVTVQEFMDNLEKETEGVLRQEMSKDLEEVKQKQVYLDLDFQKKWQEEVELYRQKVAP 144  
QY 122 LRAELQEGARQKIHLEQKLSPLGEMDRARAHVDAIRTHLAPYSDEL 170  
Db |:||:|||||  
145 LGSELREGARQKIQELQEKLSPLAELRDRARTHVDAIRQAQLAPYSDDL 193

## RESULT 4

ID	APAL_PIG	STANDARD;	PRT;	265 AA.
AC	P18648;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).			
GN	APOA1.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93224154; PubMed=8468059;			
RA	Birchbauer A., Knipping G., Juritsch B., Aschauer H., Zechner R.,			
RT	"Characterization of the apolipoprotein AI and CII genes in the			
RL	domestic pig.";			
RL	Genomics 15:643-652(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver, and Brain;			
RX	MEDLINE=94125128; PubMed=8294940;			
RA	Moockel B., Zinke H., Flach R., Weiss B., Weiler-Guettler H.,			
RT	"Expression of apolipoprotein A-I in porcine brain endothelium in			
RL	vitro.";			
RL	J. Neurochem. 62:788-798(1994).			
RN	[3]			
RP	SEQUENCE OF 34-265 FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=90132667; PubMed=2105375;			
RA	Weiler-Guettler H., Sommerfeldt M., Papandriopoulou A., Mischek U.,			
RT	Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;			
RT	"Synthesis of apolipoprotein A-I in pig brain microvascular			
RL	endothelial cells.";			
RL	J. Neurochem. 54:444-450(1990).			
RN	[4]			
RP	SEQUENCE OF 105-265 FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=93154581; PubMed=8428656;			
RA	Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;			
RT	"Sequences and expression of the porcine apolipoprotein A-I and C-III			
RL	mRNAs.";			
RL	Gene 123:173-179(1993).			
RN	[5]			
RP	SEQUENCE OF 25-265.			
RA	Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,			
RA	Jackson K., Gustavsson I., Rapacz J.;			
RL	Submitted (OCT-1995) to the SWISS-PROT data bank.			
RN	[6]			
RP	SEQUENCE OF 25-34.			
RX	MEDLINE=76184721; PubMed=178359;			
RA	Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;			
RT	"Characterization of the plasma lipoproteins and apoproteins of the			
RL	Erythrocebus patas monkey.";			
RL	Biochemistry 15:1928-1933(1976).			
CC	-1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING			





FT	CONFLICT	18	18	A -> R (IN REF. 2; CAA29858).
FT	CONFLICT	44	44	MISSING (IN REF. 3).
FT	CONFLICT	45	45	V -> I (IN REF. 2; CAA29858).
FT	CONFLICT	107	107	E -> Q (IN REF. 3).
FT	CONFLICT	123	123	Y -> F (IN REF. 2; CAA29857).
FT	CONFLICT	147	147	A -> V (IN REF. 2; CAA29858 AND REF. 3).
FT	CONFLICT	150	150	R -> G (IN REF. 2; CAA29858 AND CAA29857).
FT	CONFLICT	191	191	N -> Q (IN REF. 3).
FT	CONFLICT	195	195	MISSING (IN REF. 2; CAA29858 AND CAA29857).
FT	CONFLICT	211	211	S -> K (IN REF. 2; CAA29858 AND CAA29857).
FT	CONFLICT	255	256	VL -> LV (IN REF. 3).
FT	CONFLICT	256	256	L -> V (IN REF. 2; CAA29858 AND CAA29857).
SQ	SEQUENCE	266 AA;	30591 MW;	OFF6DB386497C7D2 CRC64;

Query Match	80.3%;	Score 711;	DB 1;	Length 266;
Best Local Similarity	78.7%;	Pred. No. 7.6e-40;		
Matches 133;	Conservative 19;	Mismatches 17;	Indels 0;	Gaps 0;

```

0Y      2  EPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSDVSTFESKLR  61
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      25  DEPRSSWDKIKIDFATVYVDVTVKDSGREYVAQFEASAFGKQLNLKLDNWDSDLSSTVSKLQ  84

```

QY	62	EQLGPAVQEEFWDNLEKETEGLRQEMSKDLEEVAKVQPYLDDFOKKQEEMELYRQKYE	121
		:	
Db	85	EQLGPAVQEEFWDNLEKETEGLRQEMSKDLEEVAKVQPYLDDFOKKQEEMERYRQKYE	144

QY 122 LRAELQEGAROKLTHELOEKLSPLEEEMDRARAHDVALRTHLAPEYDEL 170  
| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||  
Db 145 LGAELEBESAROKLTLELOEKLSPLAEEFLRDSARTHVDTLRTKLAPEYSNEL 193

RESULT	6
ID	APAI_BOVIN
AC	p15497;
DT	01-APR-1990 (Rel. 14, Created)
DT	01-MAY-1992 (Rel. 22, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Apolipoprotein A-I precursor (Apo-AI).
GN	APoAI.
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;

RP SEQUENCE FROM N. A.  
RX MEDLINE=90348478; PubMed=2117227;  
RA O'Huigin C., Chan L., Li W.H.;  
RT "Cloning and sequencing of bovine apolipoprotein A-I cDNA and  
RT molecular evolution of apolipoproteins A-I and B-100. ";  
RL Mol. Biol. Evol. 7:327-339(1990).  
RN [2]

RP SEQUENCE OF 19-265.  
RA MEDLINE=92153895; Pubmed=1739745;  
RX Sparrow D.A., Lee B.R., Laplaud M.P., Auboiron S., Bauchart D.,  
RA Chapman J.M., Gotto A.M. Jr., Yang C.Y., Sparrow J.T.;  
RT "Plasma lipid transport in the pre-ruminant calf, Bos spp: primary  
RL structure of bovine apolipoprotein A-I."; *Biochim. Biophys. Acta* 1123:145-150(1992).  
RN [3]

RP SEQUENCE OF 25-70.  
RX MEDLINE=90147795; PubMed=2105728;  
RA Auboiron S., Sparrow D.A., Beaubatie L., Bauchart D., Sparrow J.T.,  
RA Laplaud M.P., Chapman J.M.;  
RT "Characterization and amino-terminal sequence of apolipoprotein AI  
RT from plasma high density lipoproteins in the pre-ruminant calf, *Eos*  
RT spp.";

RL Biochem. Biophys. Res. Commun. 166:833-839(1990).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR	EMBL; M35870; AAA30381.1; -.			
DR	PIR; A34649; A34649.			
DR	HSSP; P02647; 1ODR.			
DR	InterPro; IPR000074; Apolipoprotein.			
DR	pfam; PF01442; Apolipoprotein; 1.			
KW	Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.			
FT	SIGNAL	1	18	
FT	PROPEP	19	24	
FT	CHAIN	25	265	APOLIPOPROTEIN A-I.
FT	DOMAIN	67	265	10 X APPROXIMATE TANDEM REPEATS.
FT	REPEAT	67	88	1.
FT	REPEAT	89	110	2.
FT	REPEAT	111	121	3 (HALF-LENGTH).
FT	REPEAT	122	143	4.
FT	REPEAT	144	165	5.
FT	REPEAT	166	187	6.
FT	REPEAT	188	209	7.
FT	REPEAT	210	230	8.
FT	REPEAT	231	241	9 (HALF-LENGTH).
FT	REPEAT	242	265	10.
FT	CONFLICT	185	186	QL -> HV (IN REF. 2).
SQ	SEQUENCE	265 AA;	30276 MW;	06A2681EA2ABA50F CRC64;

Query Match	79.4%;	Score 703;	DB 1;	Length 265;
Best Local Similarity	77.5%;	Pred. No. 2.5e-39;		
Matches 131;	Conservative 17;	Mismatches 21;	Indels 0;	Gaps 0;

```
OY      2 EPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGQLNLKLNDMWDVTSTFSKLR 61
        : ||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| : || ||| |
Db      25 DDPSQSWDRVKDFATVYVEALDKDSGRDYVAQFEASALGQLNLKLNDMDTLASTLSKYR 84
```

QY	62	EÖLGPVTOEFWMDNEKETEG	LGRQEMSKDLEEVKAKVQ	RYLDDFOFKWQEMEL	YRÖKVER	121
					:	
Db	85	EÖLGPVTOEFWMDNEKETAS	LGRQEMHKDLEEVKÖKVQ	RYLDDFOFKWHEVEI	YRÖKVAR	144

OY	122	LRAELQEGAROKLHETQEKLSP	LGEEEMDRARAHDALRTHLAPYSDEL	170
		:           :           :	:           :	
Dd	145	LGEEFREGARQKVGELQDKLSPLAQL	ELRDRAHAHVETLRQOLAPYSDDL	193

RESULT	7		
APAL_MOUSE			
ID	APAL_MOUSE	STANDARD;	PRT; 264 AA.
AC	000623;		
DT	01-APR-1993	(Rel. 25, Created)	
DT	01-APR-1993	(Rel. 25, Last sequence update)	
DT	30-MAY-2000	(Rel. 39, Last annotation update)	
DE	Apolipoprotein A-I precursor (Apo-AI).		
GN	APOA1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
OX	NCBI_TaxID=10090;		
RN	[1]		



```

RP      SEQUENCE FROM N.A.
RX      MEDLINE=92281682; Pubmed=1596360;
RA      Stoffel W., Mueller R., Binczek E., Hofmann K.;
RT      "Mouse apolipoprotein AI. cDNA-derived primary structure, gene
RL      organisation and complete nucleotide sequence.";
RN      Biol. Chem. Hoppe-Seyler 373:187-193(1992).
[2]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=93122774; Pubmed=1478650;
RA      Januzzi J.L., Azrolan N., O'Connell A., Aalto-Setälä K., Breslow J.L.;
RT      "Characterization of the mouse apolipoprotein ApoA-1/ApoC-3 gene
RL      locus: genomic, mRNA, and protein sequences with comparisons to other
RN      species.";
RX      Genomics 14:1081-1088(1992).
CC      -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC      CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC      CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC      THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC      -1- SUBCELLULAR LOCATION: Extracellular.
CC      -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC      CHYLOMICRONS.
CC      -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X64262; CAA45560.1; -.
DR      EMBL; X64263; CAA45561.1; -.
DR      EMBL; L04149; -; NOT_ANNOTATED_CDS.
DR      EMBL; L04151; -; NOT_ANNOTATED_CDS.
DR      PIR; S22420; S22420.
DR      PIR; A44364; A44364.
DR      HSSP; P02647; LAV1.
DR      SWISS-2DPAGE; Q00623; MOUSE.
DR      MGD; MGI:88049; ApoA1.
DR      InterPro; IPR000074; Apolipoprotein.
DR      Pfam; PF01442; Apolipoprotein; 1.
KW      Plasma; lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT      SIGNAL          1      18      BY SIMILARITY.
FT      PROPEP          19      24      BY SIMILARITY.
FT      CHAIN           25      264      APOLIPOPROTEIN A-1.
FT      DOMAIN          67      264      10 X APPROXIMATE TANDEM REPEATS.
FT      REPEAT          67      88      1.
FT      REPEAT          89      110      2.
FT      REPEAT          111     121      3 (HALF-LENGTH).
FT      REPEAT          111     121      4.
FT      REPEAT          122     143      5.
FT      REPEAT          144     165      6.
FT      REPEAT          166     187      7 (INCOMPLETE).
FT      REPEAT          188     207      8.
FT      REPEAT          208     229      9 (HALF-LENGTH).
FT      REPEAT          230     240      10.
FT      REPEAT          241     264
FT      SEQUENCE        264 AA; 30587 MW; C453FF2019634AAC CRC64;

Query Match          74.0%; Score 655; DB 1; Length 264;
Best Local Similarity 71.0%; Pred. No. 3.3e-36;
Matches 120; Conservative 28; Mismatches 21; Indels 0; Gaps 0
QY      2  EPPQSPWDRVYKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNDSVTSTFSKLR 61
      :  ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      25  DEPQSQWDKXVDFAENVYVDVAVKDSGRDYVSQFESSLSGQQLNLNLLENWDTLGSTVSQLO 84
QY      62  EQLGPTVQEFWNLKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWEEMELLYRQKVAP 121
      | : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      85  ERLGPLTRDFWNLEKETDWNVRQEMNKDLEEVKQKQVYPYLDEFQKKWKEDVELLYRQKVAP 144
QY      122 LRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDEL 170

```

```

DB      145  LGAELQESARQKLQELQRLSPVAEEFRDRMKRTHVDSLRTQLAPHSEQM 193
      1 11111 11111 111 :111: 1 111 1 111:11 111:::
RESULT 8
AP01_TUPGB STANDARD; PRT; 265 AA.
ID AP01_TUPGB AC 018759;
15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Apolipoprotein A-I precursor (Apo-AI).
AP01.
Tupaia glis belangeri (Common tree shrew).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
NCBI_TaxID=9396;
[1]
SEQUENCE FROM N.A.
TISUFE=LIVER;
Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;
Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
-1- SUBCELLULAR LOCATION: Extracellular.
-1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CHYLOMICRONS (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC
EMBL; AF005638; AAB82326.1; -.
DR HSSP; P02647; 1A.V1.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; signal.
FT SIGNAL; 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 265 APOLIPOPROTEIN A-I.
FT DOMAIN 67 265 10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (HALF-LENGTH).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (HALF-LENGTH).
FT REPEAT 243 265 10.
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAEA165 CRC64;
Query Match 72.0%; Score 637; DB 1; Length 265;
Best local Similarity 68.0%; Pred. No. 4.8e-35;
Matches 115; Conservative 30; Mismatches 24; Indels 0; Gaps 0
QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDVSTFESKLR 61
: 111 1111:111 1111 :1:11:1111 1 111111111:111: 111 1:
DB 25 DEPOSSWDRVRDLANVYVDVAKESGREYVSQLEASALGKQLNLKLVNDNMTLGSTFQKVH 84
QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEFVAKAVQPYLDDFQKKWQEMELYRQKVEP 121
1 11111111111: 11111 11:1:1111:1: 1 11:11: 11111:1 1111111
DB 85 EHLGPVAQEFWEKLEKETELRREINKDLEDVDYRQKTQPFLEIQKKWQEDLEERYRQKVEP 144

```

RESULT	8			
APAL_TUPGB	ID	APAL_TUPGB	STANDARD;	PRT; 265 AA.
AC	018759;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).			
OS	ApoA1.			
OC	Tupaia glis belangeri (Common tree shrew).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Scandentia; Tupaidae; Tupaia.			
OX	NCBI_TaxID=9396;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RL	Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;			
RA	Submitted (May-1997) to the EMBL/GenBank/DBJ databases.			
CC	-I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF			
CC	CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING			
CC	CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR			
CC	THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).			
CC	-I- SUBCELLULAR LOCATION: Extracellular.			
CC	-I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN			
CC	CHYLOMICRONS (BY SIMILARITY).			
CC	-I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AF005638; AAB82326.1; -.			
DR	HSSP; P02647; IAV1.			
DR	InterPro; IPR000074; Apolipoprotein.			
DR	Pfam; PF01442; Apolipoprotein; 1.			
KW	Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; signal.			
FT	SIGNAL	1	18	BY SIMILARITY.
FT	PROPEP	19	24	BY SIMILARITY.
FT	CHAIN	25	265	APOLIPOPROTEIN A-I.
FT	DOMAIN	67	265	10 X APPROXIMATE TANDEM REPEATS.
FT	REPEAT	67	88	1.
FT	REPEAT	89	110	2.
FT	REPEAT	111	121	3 (HALF-LENGTH).
FT	REPEAT	122	143	4.
FT	REPEAT	144	165	5.
FT	REPEAT	166	187	6.
FT	REPEAT	188	209	7.
FT	REPEAT	210	231	8.
FT	REPEAT	232	242	9 (HALF-LENGTH).
FT	REPEAT	243	265	10.
SEQUENCE	265 AA;	30332 MW;	60076BC39FAEA165 CRC64;	

	Query Match	72.0%;	Score 637;	DB 1;	Length 265;	
	Best Local Similarity	68.0%;	Pred. No. 4.8e-35;			
	Matches 115;	Conservative 30;	Mismatches 24;	Indels 0;	Gaps 0;	
QY	2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGQLNLKLIDNMDSVTSTESKLR	61				
	:         :         :  :                   :   :     :					
Db	25 DEPQSSWDRVRDLANVYVDVAKESGREYVSQLEASALGQLNLKLVDMPTLIGSTFQKVH	84				
QY	62 EQLPVTQEFDNLEKETEGGLRQEMSKDLEEVAKAVQPYLLDDFOKKWQEEEMELYRQKVEP	121				
	:         ::   :   :     ::         :					
Db	85 EHLPVAOEFEWEKLEKETBELRRINKKDLEDVQRQKTQPFLLDEIQKKWQEDLEERYRQKEP	144				

DR	EMBL; M00001; AAA40/49.1; -	
DR	EMBL; X00558; CAA45224.1; -	
DR	EMBL; J02597; AAA40745.1; -	
DR	PIR; A24700; A24700.	
DR	HSSP; P02647; IAV1.	
DR	InterPro; IPR000074; Apolipoprotein.	
DR	Pfam; PF01442; Apolipoprotein; 1.	
KW	Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.	
FT	SIGNAL	1 18
FT	PROPEP	19 24
FT	CHAIN	25 259
FT	DOMAIN	67 259
FT	REPEAT	67 88
FT	REPEAT	89 110
FT	REPEAT	111 121
FT	REPEAT	122 143
FT	REPEAT	4.

DR	EMBL; U086131; AAB64381.1; -.	
DR	HSSP; P02647; 1ODR.	
DR	InterPro; IPR000074; Apolipoprotein.	
DR	Pfam; PF01442; Apolipoprotein; 1.	
KW	Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal	
FT	SIGNAL	1 18 BY SIMILARITY.
FT	PROPEP	19 24 BY SIMILARITY.
FT	CHAIN	25 264 APOLIPOPROTEIN A-I.
FT	DOMAIN	67 264 10 X APPROXIMATE TANDEM REPEATS.

```
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (HALF-LENGTH).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (HALF-LENGTH).
FT REPEAT 243 264 10.
SQ SEQUENCE 264 AA; 30659 MW; 7546F4BAF2846900 CRC64;
```

Query Match 53.2%; Score 471; DB 1; Length 264;  
Best Local Similarity 52.1%; Pred. No. 2.9e-24;  
Matches 88; Conservative 38; Mismatches 43; Indels 0; Gaps 0;

```
QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKLR 61
   :||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 25 DEQAPLDRLDLDVYLETVKASGKDAIOFEASAVGKQLDLKLDNLDLGAATAKLR 84

QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELLYRQKVEP 121
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 85 EDMAPYKKEVREMWLKDTEALRAELTKDLEEVKEKIRPFLDQFSAKWTELEQYRQRLAP 144

QY 122 LRAELQEGAROKLHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDEL 170
   : ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 145 VAELKELTKQKVELMQOKLTPVAEEARDRLRGHVEELRKNLAPYSDEL 193
```

RESULT 11  
APAL\_CHICK STANDARD; PRT; 264 AA.

```
AC P08250;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88049703; PubMed=3118875;
RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.;
RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and
   evolution."
RL Biochem. Biophys. Res. Commun. 148:485-492(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88152500; PubMed=3126099;
RA Ferrari S., Tarugi P., Drusiani E., Calandra S., Fregni M.;
RT "The complete sequence of chick apolipoprotein A1 mRNA and its
   expression in the developing chick."
RL Gene 60:39-46(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87222301; PubMed=3108248;
RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,
RT "Structure, evolution, and regulation of chicken apolipoprotein A-I."
RL J. Biol. Chem. 262:7058-7065(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92381402; PubMed=1512510;
RA Lamou-Fava S., Sastri R., Ferrari S., Rajavashisth T.B.,
RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene
   expression: differences between avian and mammalian apoA-I gene
   transcription control regions."
RT
```

```
RL J. Lipid Res. 33:831-842(1992).
RN [5]
RP SEQUENCE OF 25-44.
RX MEDLINE=83213468; PubMed=6406496;
RA Shackelford J.E., Leberer H.G.;
RT "Synthesis and secretion of apolipoprotein A1 by chick breast
   muscle."
RL J. Biol. Chem. 258:7175-7180(1983).
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECTININ CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17961; AAA48593.1; -
DR EMBL; M18746; AAA48594.1; -
DR EMBL; M25559; AAA48592.1; -
DR EMBL; M96012; AAA48597.1; -
DR PIR; S01453; LPCHA1.
DR PIR; JH0471; JH0471.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein, 1.
DR Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
KW SIGNAL
FT PROPEP 1 18
FT CHAIN 25 264
FT DOMAIN 67 264
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 264
FT CONFLICT 16 16 T -> I (IN REF. 4).
FT CONFLICT 148 148 E -> K (IN REF. 3).
FT SEQUENCE 264 AA; 30680 MW; 968320E81E2AC5C2 CRC64;
```

Query Match 53.2%; Score 471; DB 1; Length 264;  
Best Local Similarity 51.5%; Pred. No. 2.9e-24;  
Matches 87; Conservative 39; Mismatches 43; Indels 0; Gaps 0;

```
QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKLR 61
   :||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 25 DEQPTPLDRIDMVDVYLETVKASGKDAIOFEASAVGKQLDLKLDNLDLGAATAKLR 84

QY 62 EQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELLYRQKVEP 121
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 85 EDMAPYKKEVREMWLKDTEALRAELTKDLEEVKEKIRPFLDQFSAKWTELEQYRQRLTP 144

QY 122 LRAELQEGAROKLHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDEL 170
   : ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 145 VAELKELTKQKVELMQAKLTPVAEEARDRLRGHVEELRKNLAPYSDEL 193
```

RESULT 12  
APAL\_COTJA







```

FT REPEAT      203    224    8.
FT REPEAT      225    246    9.
FT REPEAT      247    268   10.
FT REPEAT      269    286   11.
FT REPEAT      287    308   12.
FT REPEAT      309    330   13.
FT DOMAIN      372    420    GLU/GLN-RICH.
SQ SEQUENCE     429 AA; 49876 MW; 3D458F551D0DB60C CRC64;

Query Match          24.2%; Score 214; DB 1; Length 429;
Best Local Similarity 28.4%; Pred. No. 2.4e-07;
Matches 46; Conservative 43; Mismatches 73; Indels 0; Gaps 0;

QY 9 DRVKDLATVYVDVLKDSGRDIVSOFEGSALGKQLNLKLIDNMDSVTSTFSKREQLGPVT 68
   |:| : | : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 25 DQAVATWMDYFSQLSSNAKEAVEHLQKSELTQQLNALFDQDKLGEVNTYAGDIQKKLVFA 84

QY QEFWDNLEKETEGRLROENSKDLEEVKAKVQPYLDLFQKKWEEMELYYRQKVPEPLRAELOE 128
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 85 TELHERLAKDSEKLEKEIRKELEVRARLLPHANEVSQKIGENVRELQQRLEPYTDQLRT 144

QY 129 GARQKLHELQEKLSPLGEEMRDARAHVDAI RTHLA PYSDEL 170
   : |:| : |:| : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 145 QVNTQTQELRRQLTPYAQMEREVLRENADSLQTSLRHADQL 186

RESULT 14
APAA_PAPAN ID APAA_PAPAN STANDARD; PRT; 401 AA.
AC Q28758;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV) (Fragment).
GN APOA4.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=93340170; PubMed=8101842;
RA Hixson J.E., Kammerer C.M., Molt G.E., Britten M.L., Birnbaum S.,
RA Powers P.K., Vandeberg J.L.;
RT "Baboon apolipoprotein A-IV. Identification of Lys-76-->Glu that
RT distinguishes two common isoforms and detection of length
RT polymorphisms at the carboxyl terminus.";
RL J. Biol. Chem. 268:15667-15673(1993).
CC -!- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.
CC SECRETED IN PLASMA.
CC -!- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -!- POLYMORPHISM: THERE ARE TWO COMMON APOA-IV ISOFORMS, I (SHOWN
CC HERE) AND E. THE I ISOFORM IS ASSOCIATED WITH HIGHER LEVELS OF
CC HIGH DENSITY LIPOPROTEIN-C ON A HIGH CHOLESTEROL, SATURATED FAT
CC DIET.
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

[illegible]

RL MEDLINE=87041474; PubMed=3095836;  
RA Karathanasis S.K., Oeltgen P., Haddad I.A., Antonarakis S.E.;  
RT "Structure, evolution, and polymorphisms of the human apolipoprotein  
RL A4 gene (APOA4).";  
RN Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).  
[3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86296629; PubMed=3755616;  
RA Karathanasis S.K., Yunis I.;  
RT "Structure, evolution, and tissue-specific synthesis of human  
RT apolipoprotein AIV.";  
RL Biochemistry 25:3962-3970(1986).  
[4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87250378; PubMed=3036793;  
RA Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,  
RA Gordon J.I., Taylor J.M.;  
RT "Structure and expression of the human apolipoprotein A-IV gene.";  
RL J. Biol. Chem. 262:7973-7981(1987).  
[5]  
RP SEQUENCE OF 21-396 FROM N.A.  
RX MEDLINE=86111885; PubMed=3080432;  
RA Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;  
RT "The nucleotide and derived amino acid sequence of human  
RT apolipoprotein A-IV mRNA and the close linkage of its gene to the  
RT genes of apolipoproteins A-I and C-III.";  
RL J. Biol. Chem. 261:1998-2002(1986).  
[6]  
RP SIGNAL SEQUENCE CLEAVAGE SITE.  
RX MEDLINE=84161950; PubMed=6706947;  
RA Gordon J.I., Bisgaier C.L., Sims H.F., Sachdev O.P., Glickman R.M.,  
RA Strauss A.W.;  
RT "Biosynthesis of human preapolipoprotein A-IV.";  
RL J. Biol. Chem. 259:468-474(1984).  
[7]  
RP REVIEW ON POLYMORPHISM.  
RA Lohse P., Brewer H.B. Jr.;  
RT "Genetic polymorphism of apolipoprotein A-IV.";  
RL Curr. Opin. Lipidol. 2:90-95(1991).  
[8]  
RP VARIANT A-IV\*2.  
RX MEDLINE=90277616; PubMed=2351649;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to  
RT nucleotide substitutions in the apolipoprotein A-IV gene.";  
RL J. Biol. Chem. 265:10061-10064(1990).  
[9]  
RP VARIANTS A-IV\*0 AND A-IV\*3.  
RX MEDLINE=90324273; PubMed=1973689;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for  
RT two rare variants of apolipoprotein A-IV-1.";  
RL J. Biol. Chem. 265:12734-12739(1990).  
[10]  
RP VARIANTS.  
RX MEDLINE=91310615; PubMed=1677358;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
RT 1(Thr-347->Ser), apoA-IV-0(Lys-167->Glu,Gln-360->His), and apoA-IV-  
RT 3(Glu-165->Lys).";  
RL J. Biol. Chem. 266:13513-13518(1991).  
[11]  
RP ERRATUM.  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RL J. Biol. Chem. 266:19866-19866(1991).  
[12]  
RP VARIANT MET-13.  
RX MEDLINE=92238494; PubMed=1349197;  
RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
RA Assmann G.;  
RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
RT gene are associated with changes in the concentration of apo B- and  
RT apo A-I-containing lipoproteins in a normal population.";

RL Am. J. Hum. Genet. 50:1115-1128(1992).  
RN [13]  
RP VARIANT SER-147.  
RX MEDLINE=92144647; PubMed=1737067;  
RA Tenkanen H., Koskinen P., Metso J., Baumann M., Lukka M.,  
RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
RA Manninen V., Ehnholm C.;  
RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
RT asparagine to serine substitution at residue 127.";  
RL Biochim. Biophys. Acta 1138:27-33(1992).  
[14]  
RP VARIANT A-IV\*5.  
RX MEDLINE=93138374; PubMed=1487136;  
RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
RA Ferrell R.E., Pollitzer W.S.;  
RT "Molecular basis of a unique African variant (A-IV 5) of human  
RT apolipoprotein A-IV and its significance in lipid metabolism.";  
RL Genet. Epidemiol. 9:379-388(1992).  
[15]  
RP VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
RX MEDLINE=95245341; PubMed=7728150;  
RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
RA Csaszar A.;  
RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
RT frequencies, effect on lipid levels, and sequence of two new  
RT variants.";  
RL Hum. Mutat. 5:58-65(1995).  
[16]  
RP VARIANTS FCHE, SEATTLE SER-161; LEU-178 AND GLN-264.  
RX MEDLINE=97114287; PubMed=8956036;  
RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;  
RT "Two novel apolipoprotein A-IV variants in individuals with familial  
RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
RT activity.";  
RL Hum. Mutat. 8:319-325(1996).  
CC -I- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
CC COMPONENT OF HDL AND CHYLOMICRONS.  
CC -I- SUBCELLULAR LOCATION: Extracellular.  
CC -I- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
CC SECRETED IN PLASMA.  
CC -I- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC -I- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV\*0  
CC TO APOA-IV\*7). APOA-IV\*1 IS THE MAJOR ALLELE (90%), IV\*2 IS ALSO  
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.  
CC -I- DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN  
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF  
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).  
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M14642; AAA51745.1; -;  
DR EMBL; X13629; CAA31955.1; -;  
DR EMBL; M14566; AAA51748.1; -;  
DR EMBL; J02758; AAA96731.1; -;  
DR EMBL; M13654; AAA51744.1; -;  
DR PIR; A26481; LPHUA4.  
DR PIR; A24449; A24449.  
DR PIR; A29330; A29330.  
DR PIR; A26280; A26280.

[illegible]

Search completed: September 22, 2002, 12:22:17  
Job time: 1017 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:21:36 ; Search time 111.17 Seconds  
(without alignments)  
264.542 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_25\_194  
Perfect score: 885  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....RARAHVDAALRTHLAPYSDEL 170

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvivirus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	700	79.1	264	11	Q9Z2L4	Q9Z2L4 mesocricetu
2	645.5	72.9	263	11	Q09042	Q09042 mus musculu
3	645.5	72.9	263	11	Q08855	Q08855 mus musculu
4	565	63.8	241	6	Q9TS49	Q9TS49 erinaceus e
5	559	63.2	258	11	Q09054	Q09054 rattus norv
6	559	63.2	258	11	Q08877	Q08877 rattus norv
7	247	27.9	56	6	Q02762	Q02762 ovis aries
8	217	24.5	82	6	Q29248	Q29248 sus scrofa
9	195.5	22.1	366	13	Q93601	Q93601 gallus gall
10	186	21.0	263	13	Q98TG6	Q98TG6 anguilla ja
11	175	19.8	435	11	Q01488	Q01488 mus musculu
12	171	19.3	395	11	Q91XF8	Q91XF8 mus musculu
13	168	19.0	395	11	Q9DBN0	Q9DBN0 mus musculu
14	157.5	17.8	363	4	Q9UBJ3	Q9UBJ3 homo sapien
15	154	17.4	67	4	Q9Y355	Q9Y355 homo sapien
16	153	17.3	275	13	Q9PT02	Q9PT02 oncorhynchu

17	149.5	16.9	174	13	Q9DFQ3	Q9dfq3 gilliechthys
18	146.5	16.6	313	6	Q9GLC0	Q9glc0 tupaia glis
19	146	16.5	174	13	Q98S13	Q98s13 cyprinus ca
20	144.5	16.3	244	4	Q13784	Q13784 homo sapien
21	144	16.3	367	11	Q90UH3	Q90uh3 rattus norv
22	143.5	16.2	259	13	Q98TG5	Q98tg5 anguilla ja
23	141	15.9	26	4	Q9UCT8	Q9uct8 homo sapien
24	140.5	15.9	259	13	Q98TG4	Q98tg4 anguilla ja
25	140.5	15.9	259	13	Q98TG2	Q98tg2 anguilla ja
26	140.5	15.9	368	11	Q99P64	Q99p64 mus musculu
27	139.5	15.8	259	13	Q98TG3	Q98tg3 anguilla ja
28	139.5	15.8	368	11	Q91X90	Q91x90 mus musculu
29	136	15.4	259	13	Q98TG1	Q98tg1 anguilla ja
30	125.5	14.2	317	6	Q9GLM8	Q9glm8 gorilla gor
31	124.5	14.1	317	6	Q9GJU3	Q9gju3 pan troglod
32	124.5	14.1	317	6	Q9GLM6	Q9glm6 hylobates l
33	123.5	14.0	317	6	Q9GLM7	Q9glm7 pongo pygma
34	119.5	13.5	1547	5	Q26471	Q26471 schistocerc
35	116.5	13.2	1708	5	Q9U0S6	Q9u0s6 mytilus gal
36	115	13.0	224	13	Q9PT73	Q9pt73 scophthalmu
37	114.5	12.9	1047	11	Q9EP81	Q9ep81 mus musculu
38	113	12.8	255	4	Q96IV5	Q96iv5 homo sapien
39	113	12.8	470	5	Q9XTH4	Q9xth4 caenorhabdi
40	113	12.8	1478	4	Q9BQS8	Q9bqs8 homo sapien
41	112	12.7	1138	5	Q22276	Q22276 caenorhabdi
42	112	12.7	1411	4	Q15075	Q15075 homo sapien
43	112	12.7	1935	5	Q44934	Q44934 loligo peal
44	111.5	12.6	1046	11	Q9Z0H8	Q9z0h8 mus musculu
45	111.5	12.6	2779	5	Q9W4N7	Q9w4n7 drosophila

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	264 AA.
Q9Z2L4	Q9Z2L4			
AC	Q9Z2L4			
DT	01-MAY-1999 (TREMBlrel. 10, Created)			
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)			
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
DE	APOLIPROTEIN A-I.			
GN	APOL.			
OS	Mesocricetus auratus (Golden hamster).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Mesocricetus.			
OX	NCBI_TaxID=10036;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;			
RX	MEDLINE=99061559; PubMed=9843713;			
RA	Wu J.Y.J., Reeves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;			
RT	"Zinc deficiency decreases plasma level and hepatic mRNA abundance of			
RT	apolipoprotein A-I in rats and hamsters."			
RL	Am. J. Physiol. 275:C1516-C1525(1998).			
DR	HSSP; P02647; IAV1.			
DR	EMBL; AF046919; AAC98484.1; -.			
DR	InterPro; IPR000074; Apolipoprotein.			
DR	Pfam; PF01442; Apolipoprotein; 1.			
KW	Lipoprotein.			
SQ	SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;			

Query Match 79.1%; Score 700; DB 11; Length 264;  
Best Local Similarity 76.9%; Pred. No. 1.1e-42;  
Matches 130; Conservative 21; Mismatches 18; Indels 0; Gaps 0;

QY	2	EPQSPWDRVKDLATVYVDLKDSDGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKLR 61
Db	25	DDPQTPWDRVKDFATVYVDVAVKDSGREYVSQFETSLGKQLNLNLDNWDTLGSTVGRLLQ 84
QY	62	EQLGPTQEFWDNEKETEGLRQEMSKDLEEVKAKAVQPYLDDFQKKWQEMELYRQKVEP 121

Db 85 EOLGPVTQEFWMDNLEKETEMLRREMNKDLFEVKAQVQPYLDQFQTKWQEEVALYRQKMEP 144  
QY 122 LRAELQEGAROKLHELOEKLSPLEGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 145 LGAELRDGAROKLQELQGLSLPLGEDLRDRMRHVDALRTKMTPTSDQM 193

## RESULT 2

ID 009042 PRELIMINARY; PRT; 263 AA.

AC 009042; 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ICR AND BALB/C, AND C3H; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79575; AAB58427.1; -  
DR EMBL; U79572; AAB58424.1; -  
DR EMBL; U79573; AAB58425.1; -  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 263 AA; 30516 MW; E245DF7483A5A0DD CRC64;

## Query Match

Best Local Similarity 72.9%; Score 645.5; DB 11; Length 263;

Matches 120; Conservative 28; Mismatches 20; Indels 1; Gaps 1;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNKLNDWDSVTSTFSKLR 61  
Db 25 DEPOSQWDKVKDFANVYVDVAKDSGRDYSQFESSSLGQQLNLNLENDWDTLGSVSQLO 84  
QY 62 EOLGPVTQEFWMDNLEKETEGLRQEMSKDLEEVKAQVQPYLDQFQTKWQEEVALYRQKMEP 121  
Db 85 ERLGPLTRDFWMDNLEKETDWRQEMNKDLEEVKQKQVQPYLDQFQTKWQEEVALYRQKMEP 144  
QY 122 LRAELQEGAROKLHELOEKLSPLEGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 145 L-AELQESAROKLQELQGLSLPLGEDLRDRMRHVDALRTKMTPTSDQM 192

## RESULT 3

ID 008855 PRELIMINARY; PRT; 263 AA.

AC 008855; 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=SPLEEN;  
RX MEDLINE=98077648; PubMed=9415807;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
gene.";  
RL Biochem. Mol. Biol. Int. 43:989-996(1997).  
DR EMBL; U79574; AAB58426.1; -.

DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 263 AA; 30544 MW; 7825DF7483A5B24A CRC64;

## Query Match

Best Local Similarity 72.9%; Score 645.5; DB 11; Length 263;

Matches 120; Conservative 28; Mismatches 20; Indels 1; Gaps 1;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNKLNDWDSVTSTFSKLR 61  
Db 25 DEPOSQWDKVKDFANVYVDVAKDSGRDYSQFESSSLGQQLNLNLENDWDTLGSVSQLO 84  
QY 62 EOLGPVTQEFWMDNLEKETEGLRQEMSKDLEEVKAQVQPYLDQFQTKWQEEVALYRQKMEP 121  
Db 85 ERLGPLTRDFWMDNLEKETDWRQEMNKDLEEVKQKQVQPYLDQFQTKWQEEVALYRQKMEP 144  
QY 122 LRAELQEGAROKLHELOEKLSPLEGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 145 L-AELQESAROKLQELQGLSLPLGEDLRDRMRHVDALRTKMTPTSDQM 192

## RESULT 4

ID 09TS49 PRELIMINARY; PRT; 241 AA.

AC 09TS49; 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I, APOA-I=CHOLESTEROL TRANSPORTER.  
OS Erinaceus europaeus (Western European hedgehog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.  
OX NCBI\_TaxID=9365;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95294458; PubMed=7775860;  
RA Sparrow D.A., Laplaud P.M., Saboureau M., Zhou G., Dolphin P.J.,  
RA Gatto A.M., Jr., Sparrow J.T.;  
RT "Plasma lipid transport in the hedgehog: partial characterization of  
RT structure and function of apolipoprotein A-I.";  
RL J. Lipid Res. 36:485-495(1995).  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
SQ SEQUENCE 241 AA; 27630 MW; 2EF00F2B69210535 CRC64;

## Query Match

Best Local Similarity 63.8%; Score 565; DB 6; Length 241;

Matches 107; Conservative 22; Mismatches 36; Indels 0; Gaps 0;

QY 5 QSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNKLNDWDSVTSTFSKLRQOL 64  
Db 4 KSYWDQIKDMLTVYVDVAKDSGRDYSQFESSSLGQQLNLNLENDWDTLGSVSQLO 63  
QY 65 GPVTOEFWMDNLEKETEGLRQEMSKDLEEVKAQVQPYLDQFQTKWQEEVALYRQKMEP 124  
Db 64 KPIAMFEFQGNLEKETEGLRQVTSKDLLELVKEKQVQPYLDQFQTKWQEEVALYRQKMEP 123  
QY 125 ELQEGAROKLHELOEKLSPLEGEEMDRARAHVDALRTHLAPYSDE 169  
Db 124 EWRQEGAROKAQELQKAGELGQGHDRVRTHVDALRTDLAPYGEF 168

## RESULT 5

ID 009054 PRELIMINARY; PRT; 258 AA.

AC 009054; 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)

```

DT      01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE      APOLIPROTEIN A-I.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=WKY, AND SHRSF; TISSUE=SPLEEN;
RA      Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;
RL      Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U79578; AAB58430.1; -.
DR      EMBL; U79577; AAB58429.1; -.
DR      HSSP; P02647; 1A.V1.
DR      InterPro; IPR000074; Apolipoprotein.
DR      Pfam; PF01442; Apolipoprotein; 1.
KW      Lipoprotein.
SQ      SEQUENCE 258 AA; 29918 MW; 093E6EF2E629CDC8 CRC64;

```

Query Match	63.2%;	Score 559;	DB 11;	Length 258;
Best Local Similarity	63.5%;	Pred. NO. 1.3e-32;		
Matches 108;	Conservative 21;	Mismatches 35;	Indels 6;	Gaps 22;

QY	1	DEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEGSALGQNLKLLDNWDSVTSTFSKL	60
		:             :    :	
Db	25	DEPSQ--WDRKXDFATVYVDVAKDSGRDYSQFESSTLGQNLNLLDNWDITGSTVGR L	82
QY	61	REQLGPVTQEEFMDNLEKETEGLRQESKDL EEVKAQVOPYLDDFOKKWOEEMELYKQVE	120
		:             :          :       :    :       :	
Db	83	QEOLGVPVQDFEYANLLEKETDWP RNEMKNKDL ENVKQKMQPHIDEQOEKWNEEVEAYRQKLE	142
QY	121	PLRAELOEGAROKLHELQEKLSPLGEEMDRARAHVDA LRTHLAPYSDEL	170
		: :    :   :          :            : :	
Db	143	PLATELHKNAK----EMQRHLKVVAEEFRDRMRVNADALRAKFGLYSDQM	188

RESULT	6		
008877			
ID	008877	PRELIMINARY;	PRT; 258 AA.
AC	008877;		
DT	01-JUL-1997	(TREMBLrel. 04, Created)	
DT	01-JUL-1997	(TREMBLrel. 04, last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)	
DE	APOLIPROTEIN A-I.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SHR; TISSUE=SPLEEN;		
RX	MEDLINE=98077648; PubMed=9415807;		
RA	Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;		
RT	"Repetitive elements in the third intron of murine apolipoprotein A-I		
RT	gene.";		
RL	Biochem. Mol. Biol. Int. 43:989-996(1997).		
DR	EMBL: U79576; AAB58428.1; -.		
DR	HSSP: P02647; IAVI.		
DR	InterPro: IPR000074; Apolipoprotein.		
DR	Pfam: PF01442; Apolipoprotein; 1.		
KM	lipoprotein.		
SO	SEQUENCE 258 AA; 29831 MW; 093FEB582E629CDC8 CRC64;		

```

Query Match          63.2%; Score 559; DB 11; Length 258;
Best Local Similarity 63.5%; Pred. No. 1.3e-32;
Matches 108; Conservative 21; Mismatches 35; Indels 6; Gaps 2;

QY      1 DEPPQSPMDRVKDLATVYVVDVLKDSGRDYSQFECSALGKQLNLKLLDNWDSVTSTFSKL 60
      ||| | ||||| ||||| : ||||| ||||| ||||| ||||| : || : |
Db      25 DEPSQ--WDRVKDFEATVYVDVAKDSGRDYSQFEESSTLGKQLNLNLLDNWDTLTGSTVGRL 82

```

[illegible]

RESULT	7		
ID	002762	PRELIMINARY:	PRT; 56 AA.
AC	002762;		
DT	01-JUL-1997	(TREMBLrel. 04, Created)	
DT	01-JUL-1997	(TREMBLrel. 04, last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)	
DE	APOLIPROTEIN A1 (FRAGMENT).		
OS	Ovis aries (Sheep).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Caprinae; Ovis.		
OX	NCBI_TaxID=9940;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99098454; PubMed=9883985;		
RA	Robertson J.A., Bhattacharyya S., Ing N.H.;		
RT	"tamoxifen up-regulates oestrogen receptor-alpha, c-fos and		
RL	glycerinaldehyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium.";		
DR	J. Steroid Biochem. Mol. Biol. 67:285-292(1998).		
EMBL	U94720; AAB57840.1; -.		
HSSP	P02647; LAV1.		
KW	Lipoprotein.		
FT	NON_TER	1	1
FT	NON_TER	56	56
SEQUENCE	56 AA; 6617 MW; 2AB38E08F1E8F1BC CRC64;		

Query Match	27.9%;	Score 247;	DB 6;	Length 56;
Best Local Similarity	82.1%;	Pred. No. 4e-11;		
Matches	46;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0;

```

QY      47 LDNWDSVTSTFSKLRQGLGPVLTQEEFMDNLEKETEGLRQEMSKDLEEVKAKVQPYLD 102
      |||||: || ||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 LDNWDSLASTLSKVRQGLGPVLTQEEFMDNLEKETASLRQEMHMDLEEARQKVQPYLD 56

```

RESULT	8	
ID	Q29248	
AC	Q29248	PRELIMINARY;
DT	01-NOV-1996	(TREMBLrel. 01, Created)
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DE	APOLIPROTEIN A-I (FRAGMENT).	
OS	Sus scrofa (Pig).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	
OX	NCBI_taxid=9823;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=SMALL INTESTINE;	
RX	MEDLINE=96327607; PubMed=8672129;	
RA	Winteroe A.K., Fredholm M., Davies W.;	
RT	"Evaluation and characterization of a porcine small intestine cDNA	
RT	library."	
RL	Mamm. Genome 7:509-517(1996).	
DR	EMBL; F14858; CAA3298.1; -.	
KW	Lipoprotein.	
FT	NON_TER	82
SEQUENCE	82 AA; 9168 MW; 24625C65CBFFEDD8 CRC64;	

Query Match 24.58; Score 217; DB 6; Length 82;







## Query Match

17.8%; Score 157.5; DB 4; Length 363;

Best Local Similarity 26.1%; Pred. No. 0.0009;

Matches 43; Conservative 40; Mismatches 61; Indels 21; Gaps 4;

```
OY 24 DSGRDVVSQFEGSALGKQLNKLNDNWSVTSFESKLRQQLGPTQGFMDNLEKETEGLR 83
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 34 DKGR--VEQIHQKMKARE-PATLKDSLQDLNNMKNKFLKRLPLSGSEAPRLPQDPVGM 90
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
OY 84 QEMSKDLEEVKAKVQPYLDDEQKKWQEMELVQKVEPLRAELQEGAROKLHELQKLS 143
    ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 91 RQLQEELEEVKARLPYMAFAHVLGVNLEGLRQQLKPYTMDLMEQVALRVQELQEQLRV 150
    ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
OY 144 LGSEMR-----DRARAHVDALRTHLA-----PYSDEL 170
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 151 VGEDTKAQLLGVDEAWALLQGLQSRVHHHTGRFKELFHPYAESL 195
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
```

## RESULT 15

```
ID Q9Y355 PRELIMINARY; PRT; 67 AA.
AC Q9Y355;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE APOLIPOPROTEIN A1 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99216322; PubMed=10198255;
RA Hamidi Asl K., Liepnieks J.J., Nakamura M., Parker F., Benson M.D.;
RT "A novel apolipoprotein A-1 variant, Arg173Pro, associated with
RT cardiac and cutaneous amyloidosis.";
RL Biochem. Biophys. Res. Commun. 257:584-588(1999).
DR EMBL; AF148963; AAD34604.1; -.
DR HSSP; P02647; 1AV1.
KW Lipoprotein.
FT NON_TER 1
FT NON_TER 67
SQ SEQUENCE 67 AA; 7433 MW; 525E1FEB7BDD5AFB CRC64;
```

## Query Match

17.4%; Score 154; DB 4; Length 67;

Best Local Similarity 100.0%; Pred. No. 0.00021;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 141 LSPLGEEMRDRARAHVDALRTHLAPYSDEL 170
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 LSPLGEEMRDRARAHVDALRTHLAPYSDEL 30
```

Search completed: September 22, 2002, 12:21:36  
Job time: 1116 sec







Db 82 reqlgpvtqefwlnleketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve 141  
QY 121 PLRAELQEGAROKLHELQEKLSPLGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 142 plraelqegargklhelqeklsplgeemdrararahvdalrthlapyssel 191

RESULT 4  
AAP61079  
ID AAP61079 standard; Protein; 267 AA.  
XX  
AC AAP61079;  
XX  
DT 07-OCT-1991 (first entry)  
XX  
DE Assumed human apolipoprotein A-1 derivative gene product.  
XX  
KW Hyperlipaemia; arteriosclerosis.  
XX  
OS Homo sapiens.  
XX  
PN JP61096998-A.  
XX  
PD 15-MAY-1986.  
XX  
PF 16-OCT-1984; 84JP-0216988.  
XX  
PR 16-OCT-1984; 84JP-0216988.  
XX  
PA (MITU ) MITSUBISHI CHEM IND KK.  
XX  
DR WPI; 1986-165025/26.  
DR N-PSDB; AAN60886.  
XX  
PT Human apo: lipoprotein A-1 (deriv.) prepn. - by providing DNA  
PT fragment in cloning site downstream of expression vector promoter  
PT and introducing into host microorganism.  
XX  
PS Disclosure; Fig 2; 9pp; Japanese.  
XX  
CC The human apolipoprotein may be produced by a suitable transformed  
CC host, it is effective in treating hyperlipaemia and arteriosclerosis.  
XX  
SQ Sequence 267 AA;

Query Match 100.0%; Score 885; DB 7; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.7e-66;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdststfskl 84  
QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDEFOKKWQEMELYRQKVE 120  
Db 85 reqlgpvtqefwlnleketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve 144  
QY 121 PLRAELQEGAROKLHELQEKLSPLGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 145 plraelqegargklhelqeklsplgeemdrararahvdalrthlapyssel 194

RESULT 5  
AAP82128  
ID AAP82128 standard; Protein; 267 AA.  
XX  
AC AAP82128;  
XX  
DT 24-OCT-1990 (first entry)  
XX  
DE Entire human preproapoprotein A1.  
XX

KW human preproapoprotein A1; high density lipoprotein deficiency; ss.  
XX  
OS synthetic.  
XX  
FH Key  
FT Peptide  
FT Peptide  
FT Peptide  
FT Protein  
FT  
XX  
PN EP293357-A.  
XX  
PD 30-NOV-1988.  
XX  
PF 24-MAY-1988; 88EP-0870095.  
XX  
PR 28-MAY-1987; 87GB-0012540.  
XX  
PA (UNIO ) UCB SA.  
XX  
PI Bollen A, Gobert J, Wulfert E;  
XX  
DR WPI; 1988-339891/48.  
DR N-PSDB; AAN82064.  
XX  
PT New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures  
XX  
PS Disclosure; ; P; French.  
XX  
CC The cDNA 878bp fragment encoding preproapoprotein A1 was detected  
CC in clone PULB1609 derived from human liver cells.  
XX  
SQ Sequence 267 AA;

Query Match 100.0%; Score 885; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.7e-66;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdststfskl 84  
QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDEFOKKWQEMELYRQKVE 120  
Db 85 reqlgpvtqefwlnleketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve 144  
QY 121 PLRAELQEGAROKLHELQEKLSPLGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 145 plraelqegargklhelqeklsplgeemdrararahvdalrthlapyssel 194

RESULT 6  
AAR34032  
ID AAR34032 standard; Protein; 267 AA.  
XX  
AC AAR34032;  
XX  
DT 13-AUG-1993 (first entry)  
XX  
DE Sequence of apo A1.  
XX  
KW Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.  
XX  
OS Homo sapiens.  
XX  
PN WO9307165-A.  
XX  
PD 15-APR-1993.

```
XX 09-OCT-1992; 92WO-US08634.
XX
XX 09-OCT-1991; 91US-0774633.
XX 08-OCT-1992; 92US-0555555.
XX 28-JUN-1992; 92US-0901706.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;
XX
XX DR WPI; 1993-134378/16.
XX DR N-PSDB; AAQ40030.
XX
XX PT Polypeptide mimic of native apo B-100 and native apo A-I - useful
XX in assays for LDL and HDL in plasma samples
XX
XX PS Claim 19; Pages 105-106; 137pp; English.
XX
XX CC The inventors claim a portion of the polypeptide contg. apo B-100
XX that immunoreacts with antibodies secreted by the hybridoma MB47
XX having ATCC Accession No. 8746. Polypeptides specifically claimed
XX include residues 217-297, 216-310, 216-331, 216-352, 216-377, 1-377,
XX 205-297, 173-297, 140-297. DNA sequences encoding the polypeptides
XX are also claimed. Also claimed are a fusion polypeptide that
XX contains: (a) a first amino acid residue sequence up to 250 residues
XX in length that includes residues 120-135 of apo A-I, (b) a second
XX amino acid residue sequence up to 375 residues in length that
XX includes residues 217-297 of apo B-100 and DNA encoding it.
XX
XX SQ Sequence 267 AA;

Query Match 100.0%; Score 885; DB 14; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.7e-66;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDVYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60
Db 25 deppqspwdrvkdlatvyvdvldksgrdyvsqfegsalsgkqlnlkldnwdststfskl 84
OY 61 REQLGPTQEFWDLNLEKETEGIRQEMSKDLEEVKAKVQPYLDDFQKKWOEMELLYRQKVE 120
Db 85 reqlgptqefwdlneketeqlrgemskdleevkakvqpylddfqkkwgeemelyrqkve 144
OY 121 PLRAELQEGAROKLHELQKLSPLGEMDRARAHVDAIRTHLAPYSDEL 170
Db 145 plraelqegarqklhelqeklsplgeemdrarahvdairthlapysdel 194

RESULT 7
AAR72705
ID AAR72705 standard; Protein: 267 AA.
XX
XX AC AAR72705;
XX
XX 31-OCT-1995 (first entry)
XX
XX DE Human apo A-I including signal and propeptide sequences.
XX
XX KW Apo A-I; LDL cholesterol; low density lipoprotein; lipid.
XX
XX OS Homo sapiens.
XX
XX FH Key location/Qualifiers
XX FT Peptide 1..18
XX FT Peptide /label= presignal
XX FT Peptide 19..24
XX FT Peptide /label= propeptide
XX FT Peptide 120..135
XX FT Peptide /label= claimed
XX FT Peptide /note= "as part of fusion polypeptide"
XX FT Peptide 19..240
```

```
FT /label= claimed
FT /note= "as part of fusion polypeptide"
XX
XX PN US5408038-A.
XX
XX PD 18-APR-1995.
XX
XX PF 09-OCT-1991; 91US-0774633.
XX
XX 09-OCT-1991; 91US-0774633.
XX 08-OCT-1992; 92US-0901706.
XX 08-OCT-1992; 92US-0959946.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;
XX
XX DR WPI; 1995-161146/21.
XX DR N-PSDB; AAQ89634.
XX
XX PT New apo: lipoprotein B-100 peptide(s) and fusion peptide(s) - used
XX in assay systems for detecting LDL and HDL cholesterol levels in
XX body fluids.
XX
XX PS Claim 10; Fig 2; 41pp; English.
XX
XX CC AAA89634 and AAR72705 depict the AA sequence of human apo A-I and its
XX corresp. cDNA, including presignal residues and propeptide
XX residues, according to Seilhamer et al., DNA 3(4):309 (1984).
XX A dispersible apo A-I/B-100 fusion polypeptide is claimed which
XX contains a first AA sequence of apo A-I and that includes at
XX least AA sequence positions 120-135 (see AAR72606) and which reacts
XX with pan anti-apo A-I antibodies such as: AI-4 ATCC HB8744; AI-7
XX ATCC HB 8745; AI-9 ATCC HB 8741; AI-10 ATCC HB 9200; AI-11 ATCC
XX HB 9201; AI-12 ATCC HB 9202; AI-13 ATCC HB 9203; AI-14 ATCC HB
XX 9204; AI-18 ATCC HB 9507.
XX
XX SQ Sequence 267 AA;

Query Match 100.0%; Score 885; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.7e-66;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDVYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60
Db 25 deppqspwdrvkdlatvyvdvldksgrdyvsqfegsalsgkqlnlkldnwdststfskl 84
OY 61 REQLGPTQEFWDLNLEKETEGIRQEMSKDLEEVKAKVQPYLDDFQKKWOEMELLYRQKVE 120
Db 85 reqlgptqefwdlneketeqlrgemskdleevkakvqpylddfqkkwgeemelyrqkve 144
OY 121 PLRAELQEGAROKLHELQKLSPLGEMDRARAHVDAIRTHLAPYSDEL 170
Db 145 plraelqegarqklhelqeklsplgeemdrarahvdairthlapysdel 194

RESULT 8
AAV18675
ID AAV18675 standard; Protein: 267 AA.
XX
XX AC AAV18675;
XX
XX 09-JUL-1999 (first entry)
XX
XX DE Human apolipoprotein AI protein sequence.
XX
XX KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;
XX ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;
XX hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.
XX
XX OS Homo sapiens.
XX
```

Query Match	100.0%;	Score 885;	DB 20;	Length 267;
Best Local Similarity	100.0%;	Pred. NO. 3.7e-66;		
Matches 170;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	DEPPQSPMDRVKDLATGVVVDVLKDSGRDYVSOFGSALGKQLNLKLLDNWDSVTSTFSKL 60			
Db 25	deppqspmdrvkdlatgyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvststfskl 84			
QY 61	REQLGPVTOEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQOEEMELYRQKVE 120			
Db 85	reqlgpvtqefwdnleketeglrqemskdleevkakvqpylddfqkkqgeemelyrqkve 144			
QY 121	PLRAELQEGARQKLEHLEOEKLSPLGGEEMDRARAHVAALRTHLAPYSDEL 170			
Db 145	plraelqegarqklhelqeklspiggeemdrarahvalrthlapyssel 194			
RESULT 9				
AAB47620				
ID AAB47620	standard; Protein; 267 AA.			
XX AAB47620;				
AC				
XX 21-JAN-2002	(first entry)			
DT				
XX				
DE	Full length Apo-A1.			
XX				
KW	Apolipoprotein; Apo-A1; Apo-A-I fragment T-cell activation inhibitor;			
KW	AFT1; monocyte; IL-1; interleukin 1; TNF; tumour necrosis factor;			
KW	acute pancreatitis; Alzheimer's disease; asthma; cancer; fever;			

KW		inflammatory bowel disease; ischemia; multiple sclerosis; osteoporosis;
KW		Parkinson's disease; psoriasis; probe.
XX		
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	Binding-site	44..65
FT		/label= Helical lipid binding domain
FT	Binding-site	220..241
FT		/label= Helical lipid binding domain
FT	Domain	74..111
FT		/note= "Involved in lipoprotein-mediated cholesterol efflux from monocytes"
FT	Binding-site	149..219
FT		/label= Receptor binding domain
FT	Domain	99..120
FT		/label= Major antigenic epitope domain
FT	Domain	99..143
FT		/label= Hinged domain
FT	Domain	66..120
FT		/label= Phylogenetically conserved domain
FT	Domain	90..111
FT		/note= "Involved in lectin-cholesterol acyltransferase activity"
FT	Domain	44..65
FT		/label= Amphipathic helix
FT	Domain	66..98
FT		/label= Amphipathic helix
FT	Domain	99..120
FT		/label= Amphipathic helix
FT	Domain	121..142
FT		/label= Amphipathic helix
FT	Domain	143..164
FT		/label= Amphipathic helix
FT	Domain	165..208
FT		/label= Amphipathic helix
FT	Domain	209..219
FT		/label= Amphipathic helix
FT	Domain	220..241
FT		/label= Amphipathic helix
FT	Peptide	25..194
FT		/label= AFTI
FT	Peptide	/note= "18 kD N-terminal fragment"
FT		25..144
FT		/label= AFTI
FT	Peptide	/note= "13 kD N-terminal fragment"
FT		156..267
FT		/label= AFTI
FT		/note= "13 kD C-terminal fragment"
XX		
PN	WO200168852-A2.	
XX		
PD	20-SEP-2001.	
XX		
PF	13-MAR-2001; 2001WO-US07826.	
XX		
PR	13-MAR-2000; 2000US-189008P.	
XX		
PA	(AMGE-) AMGEN INC.	
XX		
PI	Edwards CK, Burger D, Dayer J, Kohno T;	
XX		
DR	WPI; 2001-596908/67.	
DR	N-PSTB; AAH43623.	
XX		
PT	Apo-A-I fragment T-cell activation inhibitor (AFTI) polynucleotides,	
PT	useful for treating, diagnosing, ameliorating diseases associated with	
PT	IL-1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease	
PT	and asthma -	
XX		
PS	Claim 1; Fig 1A; 132pp; English.	
CC	This sequence shows full length apolipoprotein (Apo-AI). Fragments	





CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising  
 CC the nucleic acids encoding the polypeptides and cells genetically  
 CC engineered to express them are also useful for producing the proteins.  
 CC The proteins are useful in genetic vaccination, testing and  
 CC therapy, and can be used as nutritional supplements. They may be used to  
 CC increase stem cell proliferation; to regulate haematopoiesis; and in  
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
 CC sequences of novel human secreted proteins of the invention.  
 XX  
 SQ Sequence 299 AA;

Query Match 99.2%; Score 878; DB 22; Length 299;  
 Best Local Similarity 99.4%; Pred. No. 1.6e-65;  
 Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
 Db 57 deppqspwdrvkdlatvyvdvldksgrdyvsqfegsalgkqlnlkldnwdststfskl 116  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDPQKKQWQEMELYRQKVE 120  
 Db 117 reqlgpvtqefwdnleketeglrqemskdleevkakvpyldldfpqkkqweemelyrqkve 176  
 QY 121 PLRAELQEGARQKLHELOEKISPLGEMDRARAHVDALRTHLAPYSDEL 170  
 Db 177 plraelqegarqklhelqekisplgeemdrarahvdalrthlapysgel 226

RESULT 12  
 AAW08602  
 ID AAW08602 standard; Protein; 267 AA.

XX AAW08602;

DT 04-SEP-1997 (first entry)

XX Human apolipoprotein A-1 variant "Paris" protein sequence.

KW Human; apolipoprotein A-1; variant; mutation; serum lipid; cholesterol;  
 KW high density lipoprotein; triglyceride; symptom; cardiovascular disease;  
 KW atherosclerosis; restenosis; myocardial infarct; angina; death; stroke;  
 KW cardiac decompensation; metabolic deficit.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..24  
 FT /note= "prepro peptide"  
 FT Region 172..178

FT /note= "this sequence which contains the mutated amino  
 FT Misc-difference 175 acid residue is claimed - claim 2"

FT /note= "changed from Arg residue in wild type protein,  
 FT due to a C to T transition mutation"

PN WO9637608-A1.

XX 28-NOV-1996.

XX 20-MAY-1996; 96WO-FR00747.

XX 22-MAY-1995; 95FR-0006061.

PA (INSP ) INST PASTEUR LILLE.  
 PA (RHON ) RHONE POULENC RORER SA.  
 PA (UYPA-) UNIV CURIE PARIS VI P & M.

XX Benoit P, Bruckert E, Denefle P, Duverger N, Fruchart J;  
 PI Luc G, Turping, Assmann G, Funke H;  
 XX WPI; 1997-021218/02.  
 DR N-PSDB; AAT43691.

PT New variant of human apoA-1 with Cys at position 151 - has  
 PT anti-atherogenic activity for treatment and prevention of  
 PT cardiovascular disease

PS Claim 2; Page -; 58pp; French.

CC This is the amino acid sequence of a human apolipoprotein A-1 variant  
 CC designated the "Paris" variant which has a Cys replacing the Arg residue  
 CC at position 151. The substitution is generated by a mutation of  
 CC the C nucleotide at position 523 in the wild type gene to a T residue,  
 CC changing the encoded residue from an Arg to a Cys. The gene was isolated  
 CC from a patient with an unusual pattern of serum lipids i.e. low levels of  
 CC apoA-1 and high density lipoprotein (HDL)-cholesterol and high  
 CC triglycerides, but showing symptoms of atherosclerosis. The new variant  
 CC protein has anti-atherogenic activity so is useful for treatment and  
 CC prevention of cardiovascular diseases such as atherosclerosis,  
 CC restenosis, myocardial infarct, angina, sudden death, stroke and cardiac  
 CC decompensation, or more generally any condition involving genetic or  
 CC metabolic deficit of apoA-1.  
 CC Note: this sequence is not given in the specification but is generated  
 CC from the wild type apoA-1 gene disclosed in the specification and has  
 CC the appropriate amino acid changed.

XX Sequence 267 AA;

Query Match 99.1%; Score 877; DB 18; Length 267;  
 Best Local Similarity 99.4%; Pred. No. 1.7e-65;  
 Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
 Db 25 deppqspwdrvkdlatvyvdvldksgrdyvsqfegsalgkqlnlkldnwdststfskl 84

QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDPQKKQWQEMELYRQKVE 120  
 Db 85 reqlgpvtqefwdnleketeglrqemskdleevkakvpyldldfpqkkqweemelyrqkve 144

QY 121 PLRAELQEGARQKLHELOEKISPLGEMDRARAHVDALRTHLAPYSDEL 170  
 Db 145 plraelqegarqklhelqekisplgeemdrarahvdalrthlapysdel 194

RESULT 13

AAU28184  
 ID AAU28184 standard; Protein; 244 AA.

XX AAU28184;

DT 18-DEC-2001 (first entry)

DE Novel human secretory protein, Seq ID No 353.

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
 KW fertility; analgesic; pain; antigen.

OS Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.  
XX 05-MAR-2001; 2001WO-US04942.  
XX 07-MAR-2000; 2000US-0519705.  
PR 19-MAY-2000; 2000US-0574454.  
PR 17-JUN-2000; 2000US-0596193.  
PR 14-JUL-2000; 2000US-0616847.  
PR 19-SEP-2000; 2000US-0655363.  
PR 20-OCT-2000; 2000US-0693267.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;  
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;  
XX WPI; 2001-589934/66.  
DR N-PSDB; AAS45084.

Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders -

Example 5; SEQ ID NO 353; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention.

XX Sequence 244 AA;  
SQ

Query Match 83.6%; Score 739.5; DB 22; Length 244;  
Best Local Similarity 85.9%; Pred. No. 4.6e-54;  
Matches 146; Conservative 1; Mismatches 0; Indels 23; Gaps 1;

OY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLDNMWDSVTSTFSL 60  
Db 25 depqspwdrvkdlatvyvdvldksqg-----dsvtstfsl 61  
OY 61 REQLGPTQEFWDNLEKEFEGLRQEMSKDLEFVAKAVQPYLDDFQKKWQEMELYRQKVE 120  
|||||

Db 62 reqlgpvtqefwdnleketeglrqemskdleevkavqpylddfqkkwqemelyrqkve 121  
OY 121 PLRAELQEGAROKLHELQEKLSPLGEMRDRAHAYDALRTHLAPYSD 170  
Db 122 plraelqegarqklhelqeklsplgemrdrarahvdalrthlapysdel 171  
|||||

RESULT 14  
AAO02278  
ID AAO02278 standard; Protein; 151 AA.  
XX  
AC AAO02278;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 16170.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX Homo sapiens.  
XX WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PE 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.  
XX PA  
XX PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
DR N-PSDB; AAI82209.

XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
XX  
PS Claim 20; SEQ ID NO 16170; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 151 AA;  
SQ

Query Match 82.1%; Score 727; DB 22; Length 151;  
Best Local Similarity 94.0%; Pred. No. 2.8e-53;  
Matches 140; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 22 LKDSGRDYVSQFEGSALGKQLNLKLDNMWDSVTSTFSLRQQLGPTQEFWDNLEKETEG 81  
Db 1 lkdsgrdyvsqfegcalgkqlnlkldnmwdsstfslreglqptqefwdnlektdg 60  
OY 82 LRQEMSKDLEFVAKAVQPYLDDFQKKWQEMELYRQKVEPLRAELQEGAROKLHELQEK 141  
|||||

Db 61 lrgemskdleeekakvpylddfgkqweemelyrqkeepiraeligegarqklnhelqeel 120  
QY 142 SPLGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 121 splqemldrarahvdalrthmapysdel 149

RESULT 15  
AAU29835

ID AAU29835 standard; Protein; 221 AA.

AC AAU29835;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #326.

KW Human; vaccination; gene therapy; nutritional supplement;  
stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
vaccination, testing and therapy.

PS Claim 20; Page 199; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

Sequence 221 AA;

Query Match 80.5%; Score 712; DB 22; Length 221;  
Best Local Similarity 83.6%; Pred. No. 8.1e-52;  
Matches 143; Conservative 2; Mismatches 2; Indels 24; Gaps 2;

QY 1 DEPPQSPMDRVDKDLATVYVDVL-KDSGSDYVSQFEGSALGKQLNLKLDNMDSVTSTFSK 59  
Db 38 deppqspwdrvdktlatvvyvdgltedsgk-----dsvtstfsk 74

QY 60 LREQLGPTVQEFWMDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKQOEEMELYRQKV 119  
Db 75 lrexlgpyvtqefwdnleketeglrqemskdleevkavqpylddfgkqweemelyrqkv 134  
QY 120 EPLRAELQEGAROKLHELEKLSPLGEEMDRARAHVDALRTHLAPYSDEL 170  
Db 135 eplraeligegarqklnhelqeklsplgeemdrarahvdalrthlapysdel 185

Search completed: September 22, 2002, 12:05:15  
Job time: 290 sec

Sun Sep 22 12:10:27 2002

us-09-803-918a-2\_copy\_25\_194.rag



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:06:30 ; Search time 66.91 Seconds

(without alignments)  
56.008 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_75\_113

Perfect score: 201

Sequence: 1 DSVTSTFSKLRQLGPTQGFWDNLEKETEGLRQEMSKD 39

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
NO.							
1	201	100.0	267	1	LPHUAI		apolipoprotein A-I
2	194	96.5	267	1	A26529		apolipoprotein A-I
3	194	96.5	267	2	JS0079		apolipoprotein A-I
4	180	89.6	241	2	A24998		apolipoprotein A-I
5	179	89.1	264	2	S31394		apolipoprotein A-I
6	179	89.1	265	2	A46018		apolipoprotein A-I
7	177	88.1	265	1	LPRB1B		apolipoprotein A-I
8	177	88.1	266	1	LPRB1Z		apolipoprotein A-I
9	175	87.1	231	2	JQ0704		apolipoprotein A-I
10	173	86.1	265	2	JT0672		apolipoprotein A-I
11	173	86.1	266	1	LPGDA1		apolipoprotein A-I
12	164	81.6	265	2	A56858		apolipoprotein A-I
13	143	71.1	259	2	A24700		apolipoprotein A-I
14	142	70.6	262	2	JC1237		apolipoprotein A-I
15	142	70.6	264	2	S22420		apolipoprotein A-I
16	75	37.3	264	1	LPCHA1		apolipoprotein A-I
17	73	36.3	264	2	JCS456		apolipoprotein A-I
18	72	35.8	246	2	A61448		apolipoprotein A-I
19	63	31.3	396	1	LPHUA4		apolipoprotein A-I
20	63	31.3	401	2	A47141		apolipoprotein A-I
21	63	31.3	429	2	S29565		apolipoprotein A-I
22	60	29.9	172	2	JQ0391		apolipoprotein A-I
23	59	29.4	317	2	S33450		apolipoprotein A-I
24	58	28.9	180	2	C70428		apolipoprotein E -
25	57	28.4	210	2	S64376		hypothetical prote
26	57	28.4	312	1	LPRTT		hypothetical prote
27	57	28.4	365	2	D71559		apolipoprotein E p
28	56.5	28.1	309	2	F88456		probable ABC ATPas
29	56.5	28.1	677	2	S73798		protein W03A5.4 [ti

30	56.5	28.1	1025	2	H81751	exodeoxyribonuclea
31	56	27.9	311	2	JU0036	apolipoprotein E p
32	56	27.9	365	2	B81713	recf protein, prob
33	56	27.9	986	2	T10754	cis-Golgi matrix p
34	56	27.9	1091	2	T34107	hypothetical prote
35	55.5	27.6	194	2	S76087	hypothetical prote
36	55.5	27.6	579	2	JH0820	160K golgi antigen
37	55	27.4	317	2	A28792	apolipoprotein E p
38	55	27.4	317	2	S03185	apolipoprotein E p
39	55	27.4	906	2	T45158	pre-mRNA splicing
40	55	27.4	1410	1	A57013	early endosome ant
41	54.5	27.1	377	2	D72317	hypothetical prote
42	54.5	27.1	701	2	JN0674	ubiquitin-like fus
43	54	26.9	163	2	E69133	conserved hypothet
44	54	26.9	207	2	S56209	probable membrane
45	54	26.9	2514	2	T37320	ataxia telangiecta

## ALIGNMENTS

RESULT 1  
LPHUAI  
apolipoprotein A-I precursor [validated] - human  
N/Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C/Species: Homo sapiens (man)  
C/Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C/Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010;  
6197  
R;Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A/Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A/Reference number: A90947; MUID:85026665  
A/Accession: A90947  
A/Molecule type: DNA  
A/Residues: 1-267 <SEI>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
A/Accession: B90947  
A/Molecule type: mRNA  
A/Residues: 1-267 <SE2>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R;Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988  
A/Title: Sequence and expression of Tangier apoA-I gene.  
A/Reference number: S02373; MUID:88196137  
A/Accession: S02373  
A/Molecule type: DNA  
A/Residues: 1-267 <MAK>  
A/Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
R;Shoulders, C.C.; Kornblith, A.R.; Munro, B.S.; Baralle, F.E.  
Nucleic Acids Res. 11, 2827-2837, 1983  
A/Title: Gene structure of human apolipoprotein A1.  
A/Reference number: A93465; MUID:83220822  
A/Accession: A93465  
A/Molecule type: DNA  
A/Residues: 1-267 <SHO>  
A/Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518;  
R;Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
A/Title: Isolation and characterization of the human apolipoprotein A-I gene.  
A/Reference number: A21147; MUID:84016011  
A/Accession: A21147  
A/Molecule type: DNA  
A/Residues: 1-267 <KAR>  
A/Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.  
Nucleic Acids Res. 12, 3917-3932, 1984  
A/Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundan  
A/Reference number: A93519; MUID:84221405  
A/Accession: A93519  
A/Molecule type: mRNA  
A/Residues: 1-267 <SHA>  
A/Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753

A:Accession: B93519  
 A:Molecule type: DNA  
 A:Residues: 1-24 <SH2>  
 R:Cheung, P.; Chan, L.  
 Nucleic Acids Res. 11, 3703-3715, 1983  
 A:Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
 A:Reference number: A93472; MUID:83220772  
 A:Accession: A93472  
 A:Molecule type: mRNA  
 A:Residues: 1-267 <CHE>  
 A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R:Law, S.W.; Brewer Jr., H.B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
 A:Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
 A:Reference number: A94010; MUID:84119464  
 A:Accession: A94010  
 A:Molecule type: mRNA  
 A:Residues: 1-267 <LAW>  
 A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R:Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
 A:Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
 A:Reference number: A21118; MUID:83195100  
 A:Accession: A21118  
 A:Molecule type: mRNA  
 A:Residues: 1-24 <ZAN>  
 R:Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
 Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
 A:Title: Human plasma proapoA-I: isolation and amino-terminal sequence.  
 A:Reference number: A90112; MUID:83256553  
 A:Accession: A90112  
 A:Molecule type: protein  
 A:Residues: 19-27 <BRE>  
 R:Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
 Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
 A:Title: The amino acid sequence of human ApoA-I, an apolipoprotein isolated from high density lipoproteins  
 A:Reference number: A90209; MUID:78123731  
 A:Accession: A90209  
 A:Molecule type: protein  
 A:Residues: 25-57, 'Q', 59-169, 'QQ', 172-267 <BR2>  
 R:Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
 J. Clin. Invest. 82, 803-807, 1988  
 A:Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A-I)  
 A:Reference number: A30516; MUID:88331387  
 A:Accession: A30516  
 A:Molecule type: protein  
 A:Residues: 25-56 <YUI>  
 R:Nichols, W.C.; Dwulet, F.E.; Liepnieks, J.; Benson, M.D.  
 Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
 A:Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid.  
 A:Reference number: A31582; MUID:89050104  
 A:Accession: A31582  
 A:Molecule type: protein  
 A:Residues: 25-49, 'R', 51-85, 'D', 87-107 <NIC>  
 A:Note: variant sequence from patient with familial amyloidotic polyneuropathy type IIR  
 R:Manjunath, P.; Marcel, Y.L.; Uma, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A.  
 J. Biol. Chem. 264, 16853-16857, 1989  
 A:Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
 A:Reference number: A34409; MUID:89380318  
 A:Accession: A34409  
 A:Molecule type: protein  
 A:Residues: 25-48 <MAN>  
 R:Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
 A:Title: Structural requirements of human preproapolipoprotein AI for translocation and secretion  
 A:Reference number: S02737; MUID:89149957  
 A:Accession: S02737  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-6, 'AV', 9, 'LV', 12-29 <STO>  
 A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A:Title: Transient expression of wild type and mutant human apolipoprotein AI in COS cells  
 A:Reference number: S16197; MUID:92029676  
 A:Contents: annotation; extension of studies in reference S02737  
 R:Stoffel, W.; Kruger, E.; Deutzmann, R.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
 A:Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processed in vitro  
 A:Reference number: A19913; MUID:83236195  
 A:Accession: B19913  
 A:Molecule type: protein  
 A:Residues: 1-6, 'X', 8-13, 'XXX', 17-18, 'XX', 21, 'X', 23-25, 'X', 27-29 <ST2>  
 R:Enholm, C.; Bozas, S.E.; Tenkanen, H.; Kirsbaum, L.; Metso, J.; Murphy, B.; Walke, B.  
 Biochim. Biophys. Acta 1086, 255-260, 1991  
 A:Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
 A:Reference number: A56815; MUID:92075698  
 A:Accession: A56815  
 A:Molecule type: protein  
 A:Residues: 25-31, 'P', 33 <EHN>  
 A:Experimental source: serum  
 A:Note: sequence extracted from NCBI backbone (NCBIP:69759)  
 A:Note: 32-Tip was also found  
 R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
 Biochemistry 33, 1988-1993, 1994  
 A:Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins  
 A:Reference number: A54223; MUID:94162201  
 A:Accession: A54223  
 A:Molecule type: protein  
 A:Residues: 25-39 <KUN>  
 R:Mogulievsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.; Dina, 8, 429-436, 1989  
 A:Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: purification and characterization  
 A:Reference number: I39476; MUID:89377481  
 A:Accession: I39476  
 A:Molecule type: mRNA  
 A:Residues: 19-267 <RES>  
 A:Cross-references: GB:M29068; NID:g178774; PIDN:AAA51747.1; PID:g178775  
 R:Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglun, N.; Brewer, H.B.  
 J. Biol. Chem. 263, 18530-18536, 1988  
 A:Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by the promoter region  
 A:Reference number: I39475; MUID:89054040  
 A:Accession: I39475  
 A:Molecule type: DNA  
 A:Residues: 1-14 <RE2>  
 A:Cross-references: GB:J04066; NID:g178763; PIDN:AAA51746.1; PID:g553183  
 R:Breslow, J.L.  
 Annu. Rev. Biochem. 54, 699-727, 1985  
 A:Title: Human apolipoprotein molecular biology and genetic variation.  
 A:Reference number: A90042; MUID:85278004  
 A:Contents: annotation; review of sequences, variants and gene location  
 R:Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B.  
 J. Biol. Chem. 261, 3911-3914, 1986  
 A:Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation  
 A:Reference number: A92577; MUID:86140194  
 A:Contents: annotation; acylation with palmitate  
 A:Note: an undetermined serine or threonine is acylated by fatty acid; the acylating agent is not specified  
 R:Law, S.W.; Brewer, H.B.  
 J. Biol. Chem. 260, 12810-12814, 1985  
 A:Title: Tangier disease: The complete mRNA sequence encoding for preproapo-A-I.  
 A:Reference number: I55236; MUID:86008382  
 A:Accession: I55236  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-143, 'D', 145-267 <RE3>  
 A:Cross-references: GB:M1791; NID:g178776; PIDN:AAA35545.1; PID:g178777  
 C:Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestine (HDL) in plasma.  
 C:Genetics:  
 A:Gene: GDB:APOA1  
 A:Cross-references: GDB:119684; OMIM:107680  
 A:Map position: 11q23.3-11q23.3  
 A:Introns: 15/1; 67/2  
 C:Function:  
 A:Description: participates in the reverse transport of cholesterol from tissues to the liver; stabilizes prostacyclin (PGI-2) and noncovalently binds and stabilizes prostacyclin (PGI-2)

C:Superfamily: apolipoprotein A-I  
 C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipid  
 F:1-18/Domain: signal sequence #status experimental <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-267/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 100.0%; Score 201; DB 1; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-17;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 39  
 |||||  
 Db 75 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 113

# RESULT 2

apolipoprotein A-I precursor - crab-eating macaque  
 C:Species: Macaca fascicularis (crab-eating macaque)  
 C:Date: 30-Sep-1989 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
 C:Accession: A26529; A26627; S23135; A57766  
 R:Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marotti, K.R.  
 Gene 49, 103-110, 1986  
 A:Title: The primary structure of cynomolgus monkey apolipoprotein A-I deduced from the  
 A:Reference number: A26529; MUID:87191989  
 A:Accession: A26529

A:Molecule type: mRNA  
 A:Residues: 1-267 <POL>  
 A:Cross-references: GB:M15411; NID:g342074; PIDN:AAA36834.1; PID:g342075  
 R:Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolosi,  
 Biochemistry 26, 1457-1463, 1987  
 A:Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (cyn  
 A:Reference number: A26627; MUID:87185451  
 A:Accession: A26627

A:Molecule type: protein  
 A:Residues: 25-48 <HER>  
 R:Murray, R.W.; Marotti, K.R.  
 Biochim. Biophys. Acta 1131, 207-210, 1992  
 A:Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and corres  
 A:Reference number: S23135; MUID:92305062  
 A:Accession: S23135

A:Molecule type: DNA  
 A:Residues: 1-12, 'L', 14-267 <MUR>  
 A:Cross-references: GB:M83242; NID:g342070; PIDN:AAA36832.1; PID:g342071  
 R:Sorci-Thomas, M.; Kearns, M.W.  
 J. Biol. Chem. 266, 18045-18050, 1991  
 A:Title: Transcriptional regulation of the apolipoprotein A-I gene.  
 A:Reference number: A57766; MUID:92011532  
 A:Accession: A57766

A:Molecule type: DNA  
 A:Residues: 1-10 <RES>  
 A:Cross-references: GB:M69223; NID:g342066; PIDN:AAA36831.1; PID:g553820  
 C:Comment: The precursor is synthesized in the liver and small intestine. The propeptide  
 C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HD  
 y promoting cholesterol efflux from tissues and by acting as a cofactor for the lecithin  
 C:Genetics:  
 A:Introns: 15/1; 67/2  
 C:Superfamily: apolipoprotein A-I  
 C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status predicted <PPT>  
 F:25-267/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 96.5%; Score 194; DB 1; Length 267;  
 Best Local Similarity 97.4%; Pred. No. 2.3e-16;  
 Matches 38; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 39  
 |||||  
 Db 75 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 113

RESULT 3  
 JS0079  
 apolipoprotein A-I precursor - baboon

C:Species: Papio sp. (baboon)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 13-Jun-1997  
 C:Accession: JS0079  
 R:Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; Vandenberg, J.L.  
 Gene 74, 483-490, 1988  
 A:Title: The baboon gene for apolipoprotein A-I: characterization of a cDNA clone and  
 A:Reference number: JS0079; MUID:89232739  
 A:Accession: JS0079

A:Molecule type: mRNA  
 A:Residues: 1-267 <HIX>  
 A:Experimental source: liver  
 C:Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase  
 C:Comment: This protein contains a region of repeated amino acids which form amphipat  
 C:Genetics:

A:Gene: apoA1  
 C:Superfamily: apolipoprotein A-I  
 C:Keywords: HDL; lipid binding; lipoprotein  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-267/Product: apolipoprotein A-I #status predicted <LAT>  
 F:123-144,145-166,167-188,189-210,211-232,233-254/Region: tandem repeats

Query Match 96.5%; Score 194; DB 2; Length 267;  
 Best Local Similarity 97.4%; Pred. No. 2.3e-16;  
 Matches 38; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 39  
 |||||  
 Db 75 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 113

# RESULT 4

apolipoprotein A-I - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 31-Dec-1993  
 C:Accession: A24998  
 R:Yang, C.; Yang, T.; Pownall, H.J.; Gotto Jr., A.M.  
 Eur. J. Biochem. 160, 427-431, 1986  
 A:Title: The primary structure of apolipoprotein A-I from rabbit high-density lipopro  
 A:Reference number: A24998; MUID:87030294  
 A:Accession: A24998

A:Molecule type: protein  
 A:Residues: 1-241 <YAN>  
 C:Superfamily: apolipoprotein A-I  
 C:Keywords: HDL; lipid binding; lipoprotein

Query Match 89.6%; Score 180; DB 2; Length 241;  
 Best Local Similarity 87.2%; Pred. No. 1.1e-14;  
 Matches 34; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 39  
 ||::||  
 Db 49 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 87

# RESULT 5

apolipoprotein A-I - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
 C:Accession: S31394  
 R:Moehel, B.; Flach, R.; Weiss, B.; Weller-Guettler, H.; Frey, A.; Zinke, H.; Gassen

submitted to the EMBL Data Library, November 1992  
 A:Description: Genomic organization of the porcine apolipoprotein A1 gene and study o  
 A:Reference number: S31394  
 A:Accession: S31394

A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-264 <MOE>  
A;Cross-references: EMBL:X69477; NID:g1889; PIDN:CAA49234.1; PTD:g1890  
C;Superfamily: apolipoprotein A-I

Query Match	89.18;	Score 179;	DB 2;	Length 264;
Best Local Similarity	87.28;	Pred. No. 1.6e-14;		
Matches 34;	Conservative 3;	Mismatches 2;	Indels 0;	Gaps 0;

```

Oy      1 DSVSTFSKRLRQGLPVTQEFWMDLEKETEGLRQEMSKD 39
         11: 1111111111111111111111111111111111
Db      74 DSLGSTFFKVRQGLPVTQEFWMDLEKETEGLRQEMSKD 112

```

RESULT 6  
A46018

C:Species: *Sus scrofa domestica* (domestic pig)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
C:Accession: A46018  
R:Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.  
Genomics 15, 643-652, 1993  
A:Title: Characterization of the apolipoprotein AI and CIII genes in the domestic pig.  
A:Reference number: A46018; MUID:93224154  
A:Accession: A46018  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-265 <BIR>  
A:Cross-references: GB:L00626; NID:q164358; PIDN:AAA30992.1; PID:q164359  
A:Note: sequence extracted from NCBI backbone (NCBIN:129509, NCBIPI:129511)  
C:Superfamily: apolipoprotein A-I

Query Match	89.18;	Score 179;	DB 2;	Length 265;
Best Local Similarity	87.28;	Pred. No. 1.6e-14;		
Matches 34; Conservative	3;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      1 DSVTSTSKLRQLGPTVTFEFMDNLEKETEGLRQENSKD   39
        || : ||| : | : ||||| ||||| ||||| ||||| |||||
Db       74 DSLGSFTTKVRQLGPTVTFEFMDNLEKETEARLQENSKD  112
```

## RESULT 7

**LPRB1B**

apolipoprotein A-I precursor (clone PRBA-502) - rabbit  
C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S00230; S20557  
R:Pan, T.C.; Hao, Q.L.; Yamin, T.T.; Dai, P.H.; Chen, B.S.; Chen, S.L.; Kroon, P.A.; Chae  
Eur. J. Biochem. 170, 99-104, 1987  
A:Title: Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit apolipoprotein A-I  
A:Reference number: S00230; MUID:88082866  
A:Accession: S00230  
A:Molecule type: mRNA  
A:Residues: 1-265 <PAN>  
A:Cross-references: EMBL:X06658; NID:g1461; PIDN:CAA29857.1; PID:g1462  
A:Note: the authors translated the codon AGC for residue 174 as Arg  
A:Accession: S20557  
A:Molecule type: DNA  
A:Residues: 1-17, 'R', 19-44, 'I', 46-122, 'Y', 124-146, 'V', 148-265 <PAN?>  
A:Cross-references: EMBL:X06659; NID:g1459; PIDN:CAA29858.1; PID:g1460  
C:Comment: This protein is synthesized in the small intestine.  
C:Comment: This protein is a major component of the high density lipoproteins in plasma.  
C:Genetics:  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repeat  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match	88.1%;	Score 177;	DB 1;	Length 265;
Best Local Similarity	84.6%;	Pred. No. 2.8e-14;		
Matches 33;	Conservative 5;	Mismatches 1;	Indels 0;	Gaps 0;

OY		1	DSVTSTFSKLRBQLGPTQEFWMDNEKETEGLRQEMSKD	39
			:::      :         :         :	
Dd		74	DSLSTSVSKLQEQLGPTQEFWMDNEKETEGLRQEMMKD	112

RESULT 8  
LPRB1Z

apolipoprotein A-I precursor (clone 22ap AI) - rabbit  
C/Species: *Oryctolagus cuniculus* (domestic rabbit)  
C/Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C/Accession: S06064

R;Paraskevopoulou, T.B.; Kritis, A.; Zannis, V.  
submitted to the EMBL Data Library, July 1989  
A;Reference number: S06064  
A;Accession: S06064

A;Molecule type: mRNA  
A;Residues: 1-266 <PAR>  
A;Cross-references: EMBL:X15908; NID:g1457; PIDN:CAA34024.1; PID:g1458  
C;Comment: This protein is synthesized in the small intestine.  
C;Comment: This protein is a major component of the high density lipoproteins in plasma; Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repeat  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRO>  
F;25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match	88.1%;	Score 177;	DB 1;	Length 266;
Best Local Similarity	84.6%;	Pred. No. 2.8e-14;		
Matches 33; Conservative	5;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      1 DSVTSTFSKIREQLGVPVTOEFMWNLEKETEGLRQEMSKD   39
        ||::|| |::::| |::::| |::::| |::::| |::::| |::::|
Db      74 DSLSTSYKLQEQGLGPVTOEFMWNLEKETEGLRREANMKD  112
```

RESULT  
JQ0704

apolipoprotein A-I - piq (fragment)

C:\Species: Sus scrofa domestica (domestic pig)  
C:\date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 28-Oct-1994

C;Accession: JQ0704  
R;Weiler-Guettler, H.; Sommerfeldt, M.; Papandriopoulou, A.; Mischek, U.; Bontz, D.  
J. Neurochem. 54, 444-450, 1990

A;Title: Synthesis of apolipoprotein A-1 in pig brain microvascular endothelial cells  
A;Reference number: JQ0704; MUID:90132667

A;Accession: J00704  
A;Molecule type: mRNA  
A;Residues: 1-231 <WEI>

A: Note: the authors translated the codon CAG for residue 124 as His and GAC for residue 125.  
 C: Superfamily: apolipoprotein A-I  
 C: Keywords: cholesterol metabolism; HDL; lipid binding; lipid transport; lipoprotein; lipoprotein metabolism

Query Match	87.1%;	Score 175;	DB 2;	Length 231;
Best Local Similarity	84.6%;	Pred. No. 4.1e-14;		
Matches 33; Conservative	4;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY 1 DSVNSTFSKLRQQLGPVTOQEFWMDLNEKETBGLRQEMSKD 39
    ||:||||:||||||||||||||||||||| ||:||||
DB 41 DSLGSLTFKVRQQLGPVTOEFWMDLNEKETBGLRQEMSKD 79
```

RESULT 10  
JT0672

apolipoprotein A-I - pig  
N:Alternate names: apo-A-I  
C:Species: Sus scrofa domestica (domestic pig)



```
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match                      86.1%;   Score 173;   DB 1;   Length 266;
Best Local Similarity           84.6%;   Pred. No. 8.6e-14;
Matches      33;   Conservative    4;   Mismatches    2;   Indels      0;   Gaps      0;

QY          1 DSVTSTFSKLRQLGPVTQGFWDNLEKETEGLRQEMSKD 39
||::|| :|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db          74 DSLSTVTKLRQLGPTQGFWDNLEKETEVLRQEMSKD 112

RESULT      12
A56858
apolipoprotein A-I precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 05-Jan-1996 #sequence_revision 23-Aug-1997 #text_change 13-Aug-1999
C:Accession: I45853; A56858; A34649
R:O'Huigin, C.; Chan, L.; Li, W.
Mol. Biol. Evol. 7, 327-339, 1990
A>Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolution
A:Reference number: I45853; MUID:90348478
A:Accession: I45853
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-265 <OXH>
A:Cross-references: GB:M35870; NID:g162677; PIDN:AAA30381.1; PID:g162678
R:Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Auboiiron, S.; Bauchart, D.; Chapman, M.J.; et al.
Biochim. Biophys. Acta 1123, 145-150, 1992
A>Title: Plasma lipid transport in the preterm calf, Bos spp: primary structure c
A:Reference number: A56858; MUID:92153895
A:Accession: A56858
A>Status: preliminary
A:Molecule type: protein
A:Residues: 19-184,'OL',187-265 <SPA>
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks
A>Note: sequence extracted from NCBI backbone (NCBIP:83520)
R:Auboiiron, S.; Sparrow, D.A.; Beaubatie, L.; Bauchart, D.; Sparrow, J.T.; Laplaud, P
Biochem. Biophys. Res. Commun. 166, 833-839, 1990
A>Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasm
A:Reference number: A34649; MUID:90147795
A:Accession: A34649
A:Molecule type: protein
A:Residues: 25-70 <AUB>
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks
C:Superfamily: apolipoprotein A-I
C:Keywords: lipid binding; lipoprotein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match                      81.6%;   Score 164;   DB 2;   Length 265;
Best Local Similarity           79.5%;   Pred. No. 1.1e-12;
Matches      31;   Conservative    3;   Mismatches    5;   Indels      0;   Gaps      0;

QY          1 DSVTSTFSKLRQLGPVTQGFWDNLEKETEGLRQEMSKD 39
||::|| :|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db          74 DTLASTLSKVREQLGPTQGFWDNLEKETASLRQEMHKD 112

RESULT      13
A24700
apolipoprotein A-I precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 13-Aug-1999
C:Accession: A24700; S00298; A05314
R:Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K.
J. Biol. Chem. 261, 13268-13277, 1986
A>Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-II, and
A:Reference number: A92558; MUID:87008540
A:Accession: A24700
```

Query Match	71.1%;	Score 143;	DB 2;	Length 259;
Best Local Similarity	69.2%;	Pred. No. 4e-10;		
Matches	27;	Conservative	6;	Mismatches 6; Indels 0; Gaps 0;
QY	1 DSVTSTFSKLRQLGPVTOEFWMDNEKETEGLRQEMSKD	39		
	::   :::  :           ::       ::			
Ddb	74 DTLGSTVGRLQEQLGPTQGEFWANLEKETDMLRNEMMKD	112		

Query Match	70.68;	Score 142;	DB 2;	Length 262;
Best Local Similarity	64.18;	Pred. No. 5.3e-10;		
Matches 25; Conservative	11;	Mismatches 3;	Indels 0;	Gaps 0;

```
OY      1 DSVSTFSEKLRQGLGPVTOEFWDNLKETEGTLRQENSKD   39  
       |: : | :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db     74 DTLGSTYSQLQERLGPLTRDFEWNLEKETIDWVRQEEMNKD  112
```

RESULT 15 .  
S22420  
apolipoprotein A-I precursor - mouse

	Query Match	70.6%;	Score 142;	DB 2;	Length 264;
	Best Local Similarity	64.1%;	Pred. No. 5.4e-10;		
	Matches	25;	Conservative 11;	Mismatches 3;	Indels 0; Gaps 0;
QY	1 DSVTSTFSKLR EQLGPV TQEEFWNDLEKETEG L RQEMSKD	39			
	:       :   :   :   :   :   :   :   :   :   :   :   :				
Db	74 DTLGSTVS Q LER KLGPL TRDFWMDLEKETD WVRQEMNKD	112			

Search completed: September 22, 2002, 12:06:31  
Job time: 326 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:22:19 ; Search time 35.02 Seconds  
(without alignments)  
43.120 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_75\_113

Perfect score: 201  
Sequence: 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	267	1	APAL_HUMAN P02647 homo sapien
2	194	96.5	267	1	APAL_MACFA P15568 macaca fasc
3	179	89.1	265	1	APAL_PIG P18648 sus scrofa
4	177	88.1	266	1	APAL_RABIT P09809 coryctolagus
5	173	86.1	266	1	APAL_CANFA P02648 canis famil
6	164	81.6	265	1	APAL_BOVIN P15497 bos taurus
7	143	71.1	259	1	APAL_RAT P04639 rattus norv
8	142	70.6	264	1	APAL_MOUSE P00623 mus musculu
9	137	68.2	265	1	APAL_TUPGB P18759 tupai glis
10	75	37.3	264	1	APAL_CHICK P08250 gallus gall
11	73	36.3	264	1	APAL_COTJA P32918 coturnix co
12	72	35.8	264	1	APAL_ANAPL P042296 anas platyr
13	63	31.3	396	1	APAL_HUMAN P06727 homo sapien
14	63	31.3	401	1	APAL_PAPAN P28758 papio anubi
15	63	31.3	429	1	APAL_MACFA P33621 macaca fasc
16	60	29.9	172	1	MBEB_ECOLI P13659 escherichia
17	59	29.4	317	1	APAL_PIG P18650 sus scrofa
18	58	28.9	180	1	YE76_AQUAE P67453 aquifex aeo
19	58	28.9	707	1	HS88_NEUCR P74225 neurospora
20	57	28.4	210	1	YG21_YEAST P53251 saccharomyc
21	57	28.4	312	1	APAL_RAT P02650 rattus norv
22	57	28.4	365	1	RECF_CHLTR P084077 chlamydia t
23	56.5	28.1	677	1	YD64_MYCPN P75417 mycoplasma
24	56	27.9	311	1	APAL_MOUSE P08226 mus musculu
25	56	27.9	365	1	RECF_CHLMU Q9pkw5 chlamydia m
26	56	27.9	986	1	GMI3_RAT Q62839 rattus norv
27	56	27.9	2035	1	EVPL_MOUSE Q9d952 mus musculu
28	55.5	27.6	579	1	G160_HUMAN Q08378 homo sapien
29	55	27.4	107	1	APAL_MACMU Q28502 macaca mula
30	55	27.4	107	1	APAL_SAISC Q28995 salmifri sci
31	55	27.4	317	1	APAL_MACFA P10517 macaca fasc
32	55	27.4	317	1	APAL_PAPAN P05770 papio anubi
33	55	27.4	906	1	PR01_SCHPO Q12381 schizosacch

34	54.5	27.1	475	1	FAC1_HUMAN	O75844 homo sapien
35	54	26.9	207	1	YFE6_YEAST	P43557 saccharomyc
36	54	26.9	382	1	APAL_PIG	O46409 sus scrofa
37	53.5	26.6	565	1	HEMA_IATKO	P16060 influenza a
38	53	26.4	317	1	APAL_HUMAN	P02649 homo sapien
39	53	26.4	367	1	MREC_ECOLI	P16926 escherichia
40	53	26.4	391	1	APAL_RAT	P16926 escherichia
41	53	26.4	550	1	G6PI_SCHPO	P78917 schizosacch
42	53	26.4	5327	1	ACF7_MOUSE	P02651 rattus norv
43	52.5	26.1	564	1	HEMA_IAMAP	Q9qxz0 mus musculu
44	52.5	26.1	565	1	HEMA_IAMAP	P87506 influenza a
45	52.5	26.1	629	1	GIDA_PASMU	P03454 influenza a
						P57945 pasteurella

## ALIGNMENTS

RESULT	1	STANDARD;	PRT;	267 AA.
APAL_HUMAN	APAL_HUMAN			
ID	P02647;			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).			
GN	APOL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84221405; PubMed=6328445;			
RA	Medline=84221405; PubMed=6328445;			
RA	Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,			
RA	Baralle F.E.;			
RT	"Human apolipoproteins AI, AII, CII and CIIL. cDNA sequences and mRNA			
RT	abundance.";			
RL	Nucleic Acids Res. 12:3917-3932(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8502665; PubMed=6207999;			
RA	Seihamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;			
RT	"Isolation and DNA sequence of full-length cDNA and of the entire			
RT	gene for human apolipoprotein AI -- discovery of a new genetic			
RT	polymorphism in the apo AI gene.";			
RL	DNA 3:309-317(1984).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83220822; PubMed=6406984;			
RA	Shoulders C.C., Kornblith A.R., Munro B.S., Baralle F.E.;			
RT	"Gene structure of human apolipoprotein AI.";			
RL	Nucleic Acids Res. 11:2827-2837(1983).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83220772; PubMed=6304641;			
RA	Cheung P., Chan L.;			
RT	"Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";			
RL	Nucleic Acids Res. 11:3703-3715(1983).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84119464; PubMed=6198645;			
RA	Law S.W., Brewer H.B. Jr.;			
RT	"Nucleotide sequence and the encoded amino acids of human			
RT	apolipoprotein A-I mRNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86008382; PubMed=2995392;			
RA	Law S.W., Brewer H.B. Jr.;			
RT	"Tangier disease. The complete mRNA sequence encoding for			
RT	preproapo-A-I.";			
RL	J. Biol. Chem. 260:12810-12814(1985).			
RN	[7]			

RP SEQUENCE FROM N.A.  
 RX MEDLINE=84016011; PubMed=6413973;  
 RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
 RT "Isolation and characterization of the human apolipoprotein A-I  
 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89377481; PubMed=2673706;  
 RA Mogilevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,  
 Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
 Holmquist L., Carlson L.A., Bollen A.;  
 RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
 coli: purification and biochemical characterization.";  
 RL DNA 8:429-436(1989).  
 RN [9]  
 RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
 RX MEDLINE=88196137; PubMed=3129297;  
 RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
 Zannis V.I.;  
 RT "Sequence and expression of Tangier apoA-I gene.";  
 RL Eur. J. Biochem. 173:465-471(1988).  
 RN [10]  
 RP SEQUENCE OF 118-267 FROM N.A.  
 RX MEDLINE=83091059; PubMed=6294659;  
 RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
 Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
 RT "Isolation and characterization of cDNA clones for human  
 apolipoprotein A-I.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
 RN [11]  
 RP SEQUENCE OF 19-27.  
 RX MEDLINE=83256553; PubMed=6409108;  
 RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
 Light J.A.;  
 RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
 RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
 RN [12]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=78123731; PubMed=204308;  
 RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
 Bronzert T.J.;  
 RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated  
 from high density lipoproteins.";  
 RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
 RN [13]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=75133493; PubMed=164450;  
 RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;  
 RT "The primary structure of human plasma high density apolipoprotein  
 glutamine I (APOA-I). II. The amino acid sequence and alignment of  
 cyanogen bromide fragments IV, III, and I.";  
 RL J. Biol. Chem. 250:2725-2738(1975).  
 RN [14]  
 RP SEQUENCE OF 25-56.  
 RX MEDLINE=88331387; PubMed=3047170;  
 RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
 RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
 A-I (Apo A-I). A novel function of Apo A-I.";  
 RL J. Clin. Invest. 82:803-807(1988).  
 RN [15]  
 RP SEQUENCE OF 25-48.  
 RX MEDLINE=89380318; PubMed=2506184;  
 RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
 Chapdelaine A.;  
 RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
 proteins.";  
 RL J. Biol. Chem. 264:16853-16857(1989).  
 RN [16]  
 RP SEQUENCE OF 25-43.  
 RX MEDLINE=88070603; PubMed=3120314;  
 RA Priol R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
 Pereira M.E.A.;

RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
 RT neuraminidase, to high-density lipoprotein.";  
 RL Science 238:1417-1419(1987).  
 RN [17]  
 RP SEQUENCE OF 25-42.  
 RC TISSUE-Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 RN [18]  
 RP PALMITOYLATION.  
 RX MEDLINE=86140194; PubMed=3005308;  
 RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
 RT "Human apolipoprotein A-I. Post-translational modification by fatty  
 acid acylation.";  
 RL J. Biol. Chem. 261:3911-3914(1986).  
 RN [19]  
 RP PROCESSING.  
 RX MEDLINE=83195100; PubMed=6405383;  
 RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
 Breslow J.L.;  
 RT "Intracellular and extracellular processing of human apolipoprotein  
 A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
 RN [20]  
 RP STRUCTURE BY NMR OF 190-209.  
 RX MEDLINE=96270776; PubMed=8664326;  
 RA Wang G., Treleaven W.D., Cushley R.J.;  
 RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
 presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
 and CD. Evidence for specific peptide-SDS interactions.";  
 RL Biochim. Biophys. Acta 1301:174-184(1996).  
 RN [21]  
 RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
 RX MEDLINE=98024124; PubMed=9356442;  
 RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
 RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
 lipid-bound conformation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
 RN [22]  
 RP VARIANT MILANO.  
 RX MEDLINE=83109095; PubMed=6401735;  
 RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,  
 Franceschini G., Sirtori C.R.;  
 RT "Apolipoprotein A-I-Milano. Detection of normal A-I in affected  
 subjects and evidence for a cysteine for arginine substitution in the  
 variant A-I.";  
 RL J. Biol. Chem. 258:2508-2513(1983).  
 RN [23]  
 RP VARIANT TANGIER.  
 RX MEDLINE=83300108; PubMed=6412234;  
 RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;  
 RT "Tangier disease: defective recombination of a specific Tangier  
 apolipoprotein A-I isoform (pro-apo A-I) with high density  
 lipoproteins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).  
 RN [24]  
 RP VARIANT NORWAY.  
 RX MEDLINE=84289383; PubMed=6432779;  
 RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,  
 Utermann G., Haas J., Steinmetz A., Menzel H.J., Assmann G.;  
 RT "Abnormal lecithin:cholesterol acyltransferase activation by a human  
 apolipoprotein A-I variant in which a single lysine residue is  
 deleted.";  
 RL J. Biol. Chem. 259:10063-10070(1984).  
 RN [25]  
 RP SEQUENCE OF 25-107 (VARIANT IOWA).  
 RX MEDLINE=89050104; PubMed=3142462;  
 RA Nichols W.C., Dwulet F.E., Liepnies J., Benson M.D.;  
 RT "Variant apolipoprotein AI as a major constituent of a human  
 hereditary amyloid.";



Query Match 100.0%; Score 201; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1e-17;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
Db 75 DSVTSTFSKLRQQLGPTQEFWDNLEKETEGLRQEMSKD 113

RESULT 2  
ID APAL\_MACFA STANDARD; PRT; 267 AA.  
AC P15568; P17929;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and  
OS Papio hamadryas (Hamadryas baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541, 9557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.fascicularis;  
RX MEDLINE=87191989; PubMed=3106152;  
RA Polites H.G., Melchior G.W., Castle C.K., Marotti K.R.;  
RT "The primary structure of cynomolgus monkey apolipoprotein A-I  
RT deduced from the cDNA sequence: comparison to the human sequence.";  
RL Gene 49:103-110(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.fascicularis;  
RX MEDLINE=92305062; PubMed=1610902;  
RA Murray R.W., Marotti K.R.;  
RT "Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene  
RT and corresponding flanking regions.";  
RL Biochim. Biophys. Acta 1131:207-210(1992).  
RN [3]  
RP SEQUENCE OF 25-48.  
RC SPECIES=M.fascicularis;  
RX MEDLINE=87185451; PubMed=3105581;  
RA Herbert P.N., Bauserman L.L., Lynch K.M., Saritelli A.L.,  
RA Kantor M.A., Nicolosi R.J., Shulman R.S.;  
RT "Homologues of the human C and A apolipoproteins in the Macaca  
RT fascicularis (cynomolgus) monkey.";  
RL Biochemistry 26:1457-1463(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.hamadryas; TISSUE=Liver;  
RX MEDLINE=89232739; PubMed=2907746;  
RA Hixson J.E., Borenstein S., Cox L.A., Rainwater D.L., Vandeberg J.L.;  
RT "The baboon gene for apolipoprotein A-I: characterization of a cDNA  
RT clone and identification of DNA polymorphisms for genetic studies of  
RT cholesterol metabolism.";  
RL Gene 74:483-490(1988).  
RN [5]  
RP SEQUENCE OF 1-10 FROM N.A.  
RC SPECIES=M.fascicularis;  
RA Sorci-Thomas M.;  
RL Submitted (OCT-1991) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M15411; AAA36834.1; -.  
DR EMBL; M83242; AAA36832.1; -.  
DR EMBL; M35634; AAA35380.1; -.  
DR EMBL; M69223; AAA36831.1; -.  
DR PIR; A26529; A26529.  
DR PIR; A26627; A26627.  
DR PIR; JS0079; JS0079.  
DR PIR; S23135; S23135.  
DR HSSP; P02647; 10DR.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 267  
FT DOMAIN 68 267 10 X APPROXIMATE TANDEM REPEATS.  
FT REPEAT 68 89 1.  
FT REPEAT 90 111 2.  
FT REPEAT 112 122 3 (HALF-LENGTH).  
FT REPEAT 123 144 4.  
FT REPEAT 145 166 5.  
FT REPEAT 167 188 6.  
FT REPEAT 189 210 7.  
FT REPEAT 211 232 8.  
FT REPEAT 233 243 9 (HALF-LENGTH).  
FT REPEAT 244 267 10.  
FT CONFLICT 267 267 L -> P (IN REF. 1).  
SQ SEQUENCE 267 AA; 30735 MW; 869955C024088E21 CRC64;

Query Match 96.5%; Score 194; DB 1; Length 267;  
Best Local Similarity 97.4%; Pred. NO. 7.3e-17;  
Matches 38; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
Db 75 DSVTSTFSKLRQQLGPTQEFWDNLEKETEGLRQEMSKD 113

RESULT 3  
ID APAL\_PIG STANDARD; PRT; 265 AA.  
AC P18648;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93224154; PubMed=8468059;  
RA Birchbauer A., Knipping G., Juritsch B., Aschauer H., Zechner R.;  
RT "Characterization of the apolipoprotein AI and CIII genes in the  
RT domestic pig.";  
RL Genomics 15:643-652(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Brain;  
RX MEDLINE=94125128; PubMed=8294940;  
RA Moeckel B., Zinke H., Flach R., Weiss B., Weiler-Guettler H.,  
RA Gassen H.;

```

RT "Expression of apolipoprotein A-I in porcine brain endothelium in
RT vitro."
RL J. Neurochem. 62:788-798(1994).
RN [3]
RP SEQUENCE OF 34-265 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90132667; PubMed=2105375;
RA Weiler-Guetler H., Sommerfeldt M., Papandriopoulou A., Mischek U.,
RA Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;
RT "Synthesis of apolipoprotein A-I in pig brain microvascular
RT endothelial cells."
RL J. Neurochem. 54:444-450(1990).
RN [4]
RP SEQUENCE OF 105-265 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93154581; PubMed=8428656;
RA Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;
RT "Sequences and expression of the porcine apolipoprotein A-I and C-III
RT mRNAs."
RL Gene 123:173-179(1993).
RN [5]
RP SEQUENCE OF 25-265.
RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,
RA Jackson K., Gustavsson I., Rapacz J.;
RL Submitted (OCT-1995) to the SWISS-PROT data bank.
RN [6]
RP SEQUENCE OF 25-34.
RX MEDLINE=76184721; PubMed=178359;
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
RT "Characterization of the plasma lipoproteins and apoproteins of the
RT Erythrocebus patas monkey."
RL Biochemistry 15:1928-1933(1976).
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE
CC LIVER.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L00626; AAA30992.1; -.
DR EMBL: X69477; CAA49234.1; -.
DR EMBL: X17057; -; NOT_ANNOTATED_CDS.
DR EMBL: X59414; CAA42050.1; -.
DR PIR: JQ0704; JQ0704.
DR PIR: A05311; A05311.
DR PIR: A46018; A46018.
DR PIR: S21830; S21830.
DR PIR: S31394; S31394.
DR HSSP: P02647; 1ODR.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 265
FT DOMAIN 67 265
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 142
FT REPEAT 144 165
FT REPEAT 166 187

```

```

FT REPEAT 188 209
FT REPEAT 210 230
FT REPEAT 231 241
FT REPEAT 242 265
FT CONFLICT 108 108
FT CONFLICT 143 143
FT CONFLICT 173 173
FT CONFLICT 180 180
FT CONFLICT 185 186
FT CONFLICT 209 209
FT CONFLICT 224 224
SQ SEQUENCE 265 AA; 30325 MW; 2C6E578318ECF69C CRC64;

Query Match          89.1%; Score 179; DB 1; Length 265;
Best Local Similarity 87.2%; Pred. No. 5e-15;
Matches 34; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 DSVTSIFSKLRQQLGPTQEFWDNLEKELEGRLQEMSKD 39
Db 74 DSLGTFTRKVRQQLGPTQEFWDNLEKELEGRLQEMSKD 112

RESULT 4
APAL_RABIT STANDARD; PRT; 266 AA.
ID APAL_RABIT
AC P09809;
DT 01-MAR-1989 (Rel. 10, Created)
DI 01-JAN-1990 (Rel. 13, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22AP AI; TISSUE=Small intestine;
RA Paraskopoulou T.B., Kritis A., Zannis V.I.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=88082866; PubMed=3121329;
RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,
RA Kroon P.A., Chao Y.S.;
RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit
RT apolipoprotein A-I is synthesized in the intestine but not in the
RT liver."
RL Eur. J. Biochem. 170:99-104(1987).
RN [3]
RP SEQUENCE OF 25-266.
RX MEDLINE=87030294; PubMed=3095115;
RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;
RT "The primary structure of apolipoprotein A-I from rabbit high-density
RT lipoprotein."
RL Eur. J. Biochem. 160:427-431(1986).
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90348478; PubMed=2117227;
RA O'Huigin C., Chan L., Li W.H.;
RT "Cloning and sequencing of bovine apolipoprotein A-I cDNA and
RT molecular evolution of apolipoproteins A-I and B-100.";
RL Mol. Biol. Evol. 7:327-339(1990).
RN [2]
RP SEQUENCE OF 19-265.
RX MEDLINE=92153895; PubMed=1739745;
RA Sparrow D.A., Lee B.R., Laplaud M.P., Auboiron S., Bauchart D.,
RA Chapman J.M., Gotto A.M. Jr., Yang C.Y., Sparrow J.T.;
RT "Plasma lipid transport in the pre-ruminant calf, Bos spp: primary
RT structure of bovine apolipoprotein A-I.";
RL Biochim. Biophys. Acta 1123:145-150(1992).
RN [3]
RP SEQUENCE OF 25-70.
RX MEDLINE=90147795; PubMed=2105728;
RA Auboiron S., Sparrow D.A., Beaubatie L., Bauchart D., Sparrow J.T.,
RA Laplaud M.P., Chapman J.M.;
RT "Characterization and amino-terminal sequence of apolipoprotein AI
RT from plasma high density lipoproteins in the pre-ruminant calf, Bos
RT spp.";
RL Biochem. Biophys. Res. Commun. 166:833-839(1990).
RN [4]
RP FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
RX CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
RX CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
RX THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
RN [5]
RP SUBCELLULAR LOCATION: Extracellular.
RX TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
RX CHYLOMICRONS.
RN [6]
RP SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
RN [7]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RX between the Swiss Institute of Bioinformatics and the EMBL outstation -
RX the European Bioinformatics Institute. There are no restrictions on its
RX use by non-profit institutions as long as its content is in no way
RX modified and this statement is not removed. Usage by and for commercial
RX entities requires a license agreement (See http://www.isb-sib.ch/announce/
RX or send an email to license@isb-sib.ch).
RN [8]
RP EMBL; M35870; AAA30381.1; -.
RX PIR; A34649; A34649.
RN [9]
RP HSSP; P02647; 1ODR.
RX InterPro; IPR000074; Apolipoprotein.
RN [10]
RP Pfam; PF01442; Apolipoprotein; 1.
RX Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 265
FT DOMAIN 67 265
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 230
FT REPEAT 231 241
FT REPEAT 242 265
FT CONFLICT 185 186
FT SEQUENCE 265 AA; 30276 MW; 06A2681EA2ABA50F CRC64;

```

[illegible]





```
DR HSSP; P02647; 1AV1.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 265 APOLIPOPROTEIN A-I.
FT DOMAIN 67 265 10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (HALF-LENGTH).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (HALF-LENGTH).
FT REPEAT 243 265 10.
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAEA165 CRC64;

Query Match 68.2%; Score 137; DB 1; Length 265;
Best Local Similarity 64.1%; Pred. No. 7.4e-10;
Matches 25; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQGLPVTQEFWNLEKETEGLRQEMSKD 39
   1:::|||||:| ||||| ||||| ||||| ||||| ||
Db 74 DTLGSTFOKVHEHLGVPVAGQEFWEKLEKETELRREINKD 112

RESULT 10
APAI_CHICK STANDARD; PRT; 264 AA.
AC P08250;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=88049703; PubMed=3118875;
RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.;
RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and
   evolution."
RL Biochem. Biophys. Res. Commun. 148:485-492(1987).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=88152500; PubMed=3126099;
RA Ferrari S., Tarugi P., Drusiani E., Calandra S., Fregni M.;
RT "The complete sequence of chick apolipoprotein AI mRNA and its
   expression in the developing chick."
RL Gene 60:39-46(1987).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE=87222301; PubMed=3108248;
RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,
RT Leberer H., Lusis A.J.;
RL "Structure, evolution, and regulation of chicken apolipoprotein A-I."
   J. Biol. Chem. 262:7058-7065(1987).
RN 14
RP SEQUENCE FROM N.A.
RX MEDLINE=92381402; PubMed=1512510;
RA Lamo-Fava S., Sastri R., Ferrari S., Rajavashisth T.B.,
RT Lusis A.J., Karathanasis S.K.;
RL "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene
   expression: differences between avian and mammalian apoA-I gene
   transcription control regions."
RT
```

```
RL J. Lipid Res. 33:831-842(1992).
RN [5]
RP SEQUENCE OF 25-44.
RX MEDLINE=83213468; PubMed=6406496;
RA Shackelford J.E., Leberer H.G.;
RT "Synthesis and secretion of apolipoprotein AI by chick breast
   muscle."
RL J. Biol. Chem. 258:7175-7180(1983).
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17961; AAA48593.1; -
DR EMBL; M18746; AAA48594.1; -
DR EMBL; M25559; AAA48592.1; -
DR EMBL; M96012; AAA48597.1; -
DR PIR; S01453; LPCHAL.
DR PIR; JH0471; JH0471.
DR HSSP; P02647; 10DR.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24 APOLIPOPROTEIN A-I.
FT CHAIN 25 264 10 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 67 264
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (HALF-LENGTH).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (HALF-LENGTH).
FT REPEAT 243 264 10.
FT CONFLICT 16 16 T->I (IN REF. 4).
FT CONFLICT 16 16 E->K (IN REF. 3).
SQ SEQUENCE 264 AA; 30680 MW; 968320E81E2AC5C2 CRC64;

Query Match 37.3%; Score 75; DB 1; Length 264;
Best Local Similarity 38.5%; Pred. No. 0.031;
Matches 15; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQGLPVTQEFWNLEKETEGLRQEMSKD 39
   1:::|||||:| ||||| ||||| ||||| ||||| ||
Db 74 DTLSAAAKLRQGLPVTQEFWNLEKETEGLRQEMSKD 112

RESULT 11
APAI_COTJJA STANDARD; PRT; 264 AA.
AC P32918;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Coturnix coturnix japonica (Japanese quail).
```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Columbidae;
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97212027; PubMed=9058967;
RA Oku H., Toda T., Nagata J., Ishikawa M., Neyazaki K., Shinjo C.,
RA Chinen I.;
RT "Apolipoprotein A-I of Japanese quail: cDNA sequence and modulation of
RT tissue expression by cholesterol feeding.";
RL Biosci. Biotechnol. Biochem. 61:286-290(1997).
RN [2]
RP SEQUENCE OF 25-60.
RX MEDLINE=93213845; PubMed=8461329;
RA Oku H., Ishikawa M., Nagata J., Toda T., Chinen I.;
RT "Lipoprotein and apoprotein profile of Japanese quail.";
RL Biochim. Biophys. Acta 1167:22-28(1993).
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF VLDL, HDL, LDL AND IN
CC CHYLOMICRONS. EXPRESSED IN A NUMBER OF TISSUES INCLUDING LIVER,
CC SMALL INTESTINE, LUNG, KIDNEY, HEART AND MUSCLE WITH HIGHEST
CC EXPRESSION IN LIVER AND SMALL INTESTINE.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D85133; BAA12729.1; -.
DR HSSP; P02647; 1GM4.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 264
FT DOMAIN 67 264
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 264
SQ SEQUENCE 264 AA; 30753 MW; 8781DE213C3F863F CRC64;

Query Match 36.3%; Score 73; DB 1; Length 264;
Best Local Similarity 38.5%; Pred. No. 0.054;
Matches 15; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39
ID APOA4_HUMAN STANDARD; PRT; 396 AA.
DB 74 DTLGAAAKLRDMPTPYREVREMLKDTESLRALTLTKD 112

```

```

RESULT 12
ID APOA4_HUMAN STANDARD; PRT; 264 AA.
AC 042296;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEKIN BREED; TISSUE=Liver;
RA Chen B., Lu X., Wang K., Xue H., Zeng W., Zhu D., Chi L.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U86131; AAB64381.1; -.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 264
FT DOMAIN 67 264
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 264
SQ SEQUENCE 264 AA; 30659 MW; 7546F4BAF2846900 CRC64;

Query Match 35.8%; Score 72; DB 1; Length 264;
Best Local Similarity 38.5%; Pred. No. 0.072;
Matches 15; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

```

```

QY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39
ID APOA4_HUMAN STANDARD; PRT; 396 AA.
DB 74 DTLGAAAKLRDMPTPYREVREMLKDTESLRALTLTKD 112

```

```

RESULT 13
ID APOA4_HUMAN STANDARD; PRT; 396 AA.
AC P06727;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV).
GN APOA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=89194198; PubMed=2930771;  
RA Yang C., Gu Z.W., Xiong W., Rosseneu M., Yang H.X., Lee B.M.,  
RA Gotto A.M. Jr., Chan L.;  
RT "The primary structure of human apolipoprotein A-IV."; *Biochim. Biophys. Acta* 1002:231-237(1989).  
RL [12]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=87041474; PubMed=3095836;  
RA Karathanasis S.K., Oeltgen P., Haddad I.A., Antonarakis S.E.;  
RT "Structure, evolution, and polymorphisms of the human apolipoprotein  
RT A4 gene (APOA4)."; *Proc. Natl. Acad. Sci. U.S.A.* 83:8457-8461(1986).  
RL [13]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=86296629; PubMed=3755616;  
RA Karathanasis S.K., Yunis I.;  
RT "Structure, evolution, and tissue-specific synthesis of human  
RT apolipoprotein AIV."; *Biochemistry* 25:3962-3970(1986).  
RL [14]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=87250378; PubMed=3036793;  
RA Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,  
RA Gordon J.I., Taylor J.M.;  
RT "Structure and expression of the human apolipoprotein A-IV gene."; *J. Biol. Chem.* 262:7973-7981(1987).  
RL [15]  
RN SEQUENCE OF 21-396 FROM N.A.  
RX MEDLINE=86111885; PubMed=3080432;  
RA Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;  
RT "The nucleotide and derived amino acid sequence of human  
RT apolipoprotein A-IV mRNA and the close linkage of its gene to the  
RT genes of apolipoproteins A-I and C-III."; *J. Biol. Chem.* 261:1998-2002(1986).  
RL [16]  
RN SIGNAL SEQUENCE CLEAVAGE SITE.  
RX MEDLINE=84161950; PubMed=6706947;  
RA Gordon J.I., Bisgaler C.L., Sims H.F., Sachdev O.P., Glickman R.M.,  
RA Strauss A.W.;  
RT "Biosynthesis of human preapolipoprotein A-IV."; *J. Biol. Chem.* 259:468-474(1984).  
RL [17]  
RN REVIEW ON POLYMORPHISM.  
RP Lohse P., Brewer H.B. Jr.;  
RA "Genetic polymorphism of apolipoprotein A-IV."; *Curr. Opin. Lipidol.* 2:90-95(1991).  
RL [18]  
RN VARIANT A-IV\*2.  
RX MEDLINE=90277616; PubMed=2351649;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to  
RT nucleotide substitutions in the apolipoprotein A-IV gene."; *J. Biol. Chem.* 265:10061-10064(1990).  
RL [19]  
RN VARIANTS A-IV\*0 AND A-IV\*3.  
RX MEDLINE=90324273; PubMed=1973689;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for  
RT two rare variants of apolipoprotein A-IV-1."; *J. Biol. Chem.* 265:12734-12739(1990).  
RL [10]  
RN VARIANTS.  
RX MEDLINE=91310615; PubMed=1677358;  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
RT 1(Thr-347-->Ser), apoA-IV-0(Lys-167-->Glu,Gln-360-->His), and apoA-IV-  
RT 3(Glu-165-->Lys)."; *J. Biol. Chem.* 266:13513-13518(1991).

RN [11]  
RP ERRATUM.  
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
RL J. Biol. Chem. 266:19866-19866(1991).  
RN [12]  
RP VARIANT MET-13.  
RX MEDLINE=92238494; PubMed=1349197;  
RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
RA Assmann G.;  
RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
RT gene are associated with changes in the concentration of apo B- and  
RT apo A-I-containing lipoproteins in a normal population."; *Am. J. Hum. Genet.* 50:1115-1128(1992).  
RL [13]  
RN VARIANT SER-147.  
RX MEDLINE=92144647; PubMed=1737067;  
RA Tenkunen H., Koskinen P., Metso J., Baumann M., Lukka M.,  
RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
RA Manninen V., Ehnholm C.;  
RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
RT asparagine to serine substitution at residue 127."; *Biochim. Biophys. Acta* 1138:27-33(1992).  
RL [14]  
RN VARIANT A-IV\*5.  
RX MEDLINE=93138374; PubMed=1487136;  
RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
RA Ferrell R.E., Pollitzer W.S.;  
RT "Molecular basis of a unique African variant (A-IV 5) of human  
RT apolipoprotein A-IV and its significance in lipid metabolism."; *Genet. Epidemiol.* 9:379-388(1992).  
RL [15]  
RN VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
RX MEDLINE=95245341; PubMed=7728150;  
RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
RA Csaszar A.;  
RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
RT frequencies, effect on lipid levels, and sequence of two new  
RT variants."; *Hum. Mutat.* 5:58-65(1995).  
RL [16]  
RN VARIANTS FCGL SEATTLE SER-161; LEU-178 AND GLN-264.  
RX MEDLINE=97114287; PubMed=8956036;  
RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;  
RT "Two novel apolipoprotein A-IV variants in individuals with familial  
RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
RT activity."; *Hum. Mutat.* 8:319-325(1996).  
RL [17]  
RN FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
CC COMPONENT OF HDL AND CHYLOMICRONS.  
CC [18]  
RN SUBCELLULAR LOCATION: Extracellular.  
CC [19]  
RN TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
CC [20]  
RN SECRETED IN PLASMA.  
CC [21]  
RN DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC [22]  
RN POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV\*0  
CC TO APOA-IV\*7). APOA-IV\*1 IS THE MAJOR ALLELE (90%), IV\*2 IS ALSO  
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.  
CC [23]  
RN DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN  
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF  
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).  
CC [24]  
RN SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC [25]  
RN This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial



entties requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; M14642; AAA51745.1; -  
DR EMBL; X13629; CAA31955.1; -  
DR EMBL; M14566; AAA51748.1; -  
DR EMBL; J02758; AAA96731.1; -  
DR EMBL; M13654; AAA51744.1; -  
DR PIR; A26481; LPHUA4.  
DR PIR; A24449; A24449.  
DR PIR; A29330; A29330.  
DR PIR; A26280; A26280.  
DR PIR; S02715; S02715.  
DR HSSP; P02649; INFO.  
DR SWISS-2DPAGE; P06727; HUMAN.  
DR MIM; 107690; -  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; lipid transport; HDL; Chylomicron; Repeat; Signal;  
KW Disease mutation; Polymorphism.  
FT SIGNAL 1 20  
FT CHAIN 21 396  
FT DOMAIN 33 330  
FT REPEAT 33 54  
FT REPEAT 60 81  
FT REPEAT 82 103  
FT REPEAT 115 136  
FT REPEAT 137 158  
FT REPEAT 159 180  
FT REPEAT 181 202  
FT REPEAT 203 224  
FT REPEAT 225 246  
FT REPEAT 247 268  
FT REPEAT 269 286  
FT REPEAT 287 308  
FT REPEAT 309 330  
FT DOMAIN 372 389  
FT VARIANT 13 13  
FT VARIANT 44 44  
FT VARIANT 147 147  
FT VARIANT 161 161

Query Match 31.3%; Score 63; DB 1; Length 396;  
Best Local Similarity 27.8%; Pred. No. 1.5;  
Matches 10; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 3 VTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSK 38  
Db 256 ISASAEFLRQRLAPLAEDVRGNLKGNTGGLQKSLAE 291

RESULT 14  
AP04\_PAPAN STANDARD; PRT; 401 AA.  
AC 028758;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Apolipoprotein A-IV precursor (Apo-AIV) (Fragment).  
GN APOA4.  
OS Papio anubis (Olive baboon).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=93340170; PubMed=8101842;

RA Hixson J.E., Kammerer C.M., Mott G.E., Britten M.L., Birnbaum S.,  
RA Powers P.K., Vandeberg J.L.;  
RT "Baboon apolipoprotein A-IV. Identification of Lys-76-->Glu that  
RT distinguishes two common isoforms and detection of length  
RT polymorphisms at the carboxyl terminus."  
RL J. Biol. Chem. 268:15667-15673(1993).  
CC -!- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
CC COMPONENT OF HDL AND CHYLOMICRONS.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
CC SECRETED IN PLASMA.  
CC -!- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC -!- POLYMORPHISM: THERE ARE TWO COMMON APOA-IV ISOFORMS, I (SHOWN  
CC HERE) AND E. THE I ISOFORM IS ASSOCIATED WITH HIGHER LEVELS OF  
CC HIGH DENSITY LIPOPROTEIN-C ON A HIGH CHOLESTEROL, SATURATED FAT  
CC DIET.  
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; U13174; AAA35379.1; -  
DR HSSP; P02649; INFO.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; lipid transport; HDL; Chylomicron; Repeat; Signal;  
KW NON\_TER 1 1  
FT SIGNAL <1 4  
FT CHAIN 5 401  
FT DOMAIN 17 314  
FT REPEAT 17 38  
FT REPEAT 44 65  
FT REPEAT 66 87  
FT REPEAT 99 120  
FT REPEAT 121 142  
FT REPEAT 143 164  
FT REPEAT 165 186  
FT REPEAT 187 208  
FT REPEAT 209 230  
FT REPEAT 231 252  
FT REPEAT 253 270  
FT REPEAT 271 292  
FT REPEAT 293 314  
FT DOMAIN 356 394  
FT VARIANT 80 80  
SQ SEQUENCE 401 AA; 46538 MW; 0A76D1284AA9837F CRC64;

Query Match 31.3%; Score 63; DB 1; Length 401;  
Best Local Similarity 27.8%; Pred. No. 1.5;  
Matches 10; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

QY 3 VTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSK 38  
Db 240 ISASAEFLRQRLAPLAEDMRGNLKGNTGGLQKSLAE 275

RESULT 15  
AP04\_MACFA STANDARD; PRT; 429 AA.  
ID AP04\_MACFA

Db 256 ISASAEELRQRLAPLAEDMRGNLRGNTGLOKSLAE 291  
Search completed: September 22, 2002, 12:22:19  
Job time: 1019 sec

AC P33621; 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Apolipoprotein A-IV precursor (Apo-AIV).  
GN APOA4.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leukocyte;  
RX MEDLINE=93192330; PubMed=8448212;  
RA Osada J., Pocovi M., Nicolosi R.J., Schaefer E.J., Ordovas J.M.;  
RT "Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III  
and A-IV genes.";  
RL Biochim. Biophys. Acta 1172:335-339(1993).  
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
COMPONENT OF HDL AND CHYLOMICRONS.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.  
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: X68361; CAA48421.1; -.  
DR PIR: S29565; S29565.  
DR PIR: S30195; S30195.  
DR HSSP: P02649; INFO.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
KW Plasma; lipid transport; HDL; Chylomicron; Repeat; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 429 BY SIMILARITY.  
FT DOMAIN 33 330 APOLIPOPROTEIN A-IV.  
FT REPEAT 33 54 13 X 22 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 60 81 1.  
FT REPEAT 82 103 2.  
FT REPEAT 115 136 3.  
FT REPEAT 137 158 4.  
FT REPEAT 159 180 5.  
FT REPEAT 181 202 6.  
FT REPEAT 203 224 7.  
FT REPEAT 225 246 8.  
FT REPEAT 247 268 9.  
FT REPEAT 269 286 10.  
FT REPEAT 287 308 11.  
FT REPEAT 309 330 12.  
FT REPEAT 372 420 13.  
FT DOMAIN 429 420 GLU/GLN-RICH.  
SQ SEQUENCE 429 AA; 49876 MW; 3D458F551D0DB60C CRC64;

Query Match 31.3%; Score 63; DB 1; Length 429;  
Best Local Similarity 27.8%; Pred. NO. 1.6;  
Matches 10; Conservative 15; Mismatches 11; Indels 0; Gaps 0;  
QY 3 VTST\*SKLRQLGPTQEFWDNLEKETEGLRQEMSK 38

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:21:39 ; Search time 111.17 Seconds  
(without alignments)  
60.689 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_75\_113  
Perfect score: 201  
Sequence: 1 DSVTSTFSKLRQLGPTQGFWDNLEKETEGLRQEMSKD 39

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: SP archaea:\*
  - 2: SP bacteria:\*
  - 3: SP fungi:\*
  - 4: SP human:\*
  - 5: SP invertebrate:\*
  - 6: SP mammal:\*
  - 7: SP mhc:\*
  - 8: SP organelle:\*
  - 9: SP phage:\*
  - 10: SP plant:\*
  - 11: SP rodent:\*
  - 12: SP virus:\*
  - 13: SP vertebrate:\*
  - 14: SP unclassified:\*
  - 15: SP rvirus:\*
  - 16: SP bacteriap:\*
  - 17: SP archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	167	83.1	56	6	002762	002762 ovis aries
2	155	77.1	264	11	092214	092214 mesocricetu
3	142	70.6	263	11	009042	009042 mus musculu
4	142	70.6	263	11	008855	008855 mus musculu
5	133	66.2	258	11	009054	009054 rattus norv
6	133	66.2	258	11	008877	008877 rattus norv
7	131	65.2	241	6	09TS49	09TS49 erinaceus e
8	64.5	32.1	329	11	09D495	09D495 mus musculu
9	63	31.3	244	4	Q13784	Q13784 homo sapien
10	60.5	30.1	562	12	067018	067018 influenza a
11	59.5	29.6	259	13	Q98TG1	Q98TG1 anguilla ja
12	58	28.9	311	6	Q95K18	Q95K18 macaca fasc
13	57	28.4	90	2	Q9LCR0	Q9LCR0 paenibacill
14	57	28.4	172	2	Q9F898	Q9F898 escherichia
15	57	28.4	172	2	069425	069425 enterobacte
16	56.5	28.1	131	12	Q67047	Q67047 influenza a

17	56.5	28.1	309	5	Q23138	Q23138 caenorhabdi
18	56.5	28.1	562	12	Q67017	Q67017 influenza a
19	56.5	28.1	562	12	Q67032	Q67032 influenza a
20	56.5	28.1	562	12	Q67048	Q67048 influenza a
21	56.5	28.1	565	12	Q67010	Q67010 influenza a
22	56.5	28.1	1025	16	Q9PLT7	Q9PLT7 chlamydia m
23	56	27.9	444	11	Q9CV95	Q9CV95 mus musculu
24	56	27.9	1091	5	Q18082	Q18082 caenorhabdi
25	55.5	27.6	194	16	Q55564	Q55564 synechocyst
26	55.5	27.6	348	16	Q98KF5	Q98KF5 rhizobium l
27	55.5	27.6	562	12	Q67016	Q67016 influenza a
28	55.5	27.6	562	12	Q67326	Q67326 influenza a
29	55.5	27.6	562	12	Q67143	Q67143 influenza a
30	55.5	27.6	562	12	Q67325	Q67325 influenza a
31	55.5	27.6	562	12	Q67327	Q67327 influenza a
32	55.5	27.6	565	12	Q82544	Q82544 influenza a
33	55.5	27.6	566	12	Q82751	Q82751 influenza v
34	55.5	27.6	566	12	Q82545	Q82545 influenza a
35	55.5	27.6	566	12	Q91808	Q91808 influenza a
36	55.5	27.6	566	12	Q91807	Q91807 influenza a
37	55.5	27.6	566	12	Q91806	Q91806 influenza a
38	55.5	27.6	1530	4	Q43241	Q43241 homo sapien
39	55	27.4	172	2	Q52254	Q52254 salmonella
40	55	27.4	270	4	Q9H738	Q9H738 homo sapien
41	55	27.4	313	6	Q9GLC0	Q9GLC0 tupaya galls
42	55	27.4	366	13	Q93601	Q93601 gallus gall
43	55	27.4	687	5	Q9N3W4	Q9N3W4 caenorhabdi
44	55	27.4	1410	4	Q14221	Q14221 homo sapien
45	55	27.4	1411	4	Q15075	Q15075 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	56 AA.
002762	002762			
AC	002762;			
DT	01-JUL-1997 (TREMBLrel. 04, Created)			
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	APOLIPROTEIN A1 (FRAGMENT).			
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_TaxID=9940;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99098454; PubMed=9883985;			
RA	Robertson J.A., Bhattacharyya S., Ing N.H.;			
RT	"Tamoxifen up-regulates oestrogen receptor-alpha, c-fos and			
RT	glyceroldehyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium."			
RL	J. Steroid Biochem. Mol. Biol. 67:285-292(1998).			
DR	EMBL: U94720; AAB57840.1; -.			
DR	HSSP; P02647; LAV1.			
KW	Lipoprotein.			
FT	NON_TER	1		
FT	NON_TER	56		
FT	NON_TER	56		
SO	SEQUENCE	56 AA;	6617 MW;	2AB38E08F1E8F1BC CRC64;

Query Match 83.1%; Score 167; DB 6; Length 56;  
Best Local Similarity 82.1%; Pred. No. 1.9e-14;  
Matches 32; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY	1	DSVTSTFSKLRQLGPTQGFWDNLEKETEGLRQEMSKD 39
Db	5	DSLASTLSKVRQLGPTQGFWDNLEKETASLRQEMHKD 43
RESULT	2	Q9Z214

ID	Q9Z2L4	PRELIMINARY;	PRT;	264 AA.
AC	Q9Z2L4;			
DT	01-MAY-1999	(TREMBLrel. 10, Created)		
DT	01-MAY-1999	(TREMBLrel. 10, last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, last annotation update)		
DE	APOLIPOPROTEIN A-I.			
GN	APOAI.			
OS	Mesocricetus auratus (Golden hamster).			
OC	Eukaryota; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Mesocricetus.			
OX	NCBI_TaxID=10036;			
RN	11			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;			
RA	MEDLINE=99061559; PubMed=9843713;			
RA	Wu J.Y.J., Reaves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;			
RT	"Zinc deficiency decreases plasma level and hepatic mRNA abundance of			
RT	apolipoprotein A-I in rats and hamsters."			
RL	Am. J. Physiol. 275:Cl516-Cl525(1998).			
DR	EMBL: AF046919; AAC98484.1; .			
DR	HSSP: P02647; 1AV1.			
DR	InterPro: IPR000074; Apolipoprotein.			
KM	Pfam: PF01442; Apolipoprotein; 1.			
DR	lipoprotein.			
SQ	SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;			

Query Match	77.1%;	Score 155;	DB 11;	Length 264;
Best Local Similarity	74.4%;	Pred. No. 3.8e-12;		
Matches 29; Conservative	6;	Mismatches 4;	Indels 0;	Gaps 0;

QY	Db
1	DSVTSTFSKRLQGLPVTQEFWMDLKEKETEGLRQEMSKD 39
74	DTLGSTVGRQEQGLPVTQEFWMDLKEKETEWLRREMNKD 112

RESULT	3	000000
009042		000000

AC 009042;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
OX NCB1\_TaxID=10090;  
RN 111

RP SEQUENCE FROM N.A.  
RC STRAIN=ICR AND BALB/C, AND C3H; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79575; AAB58427.1; -.  
DR EMBL; U79572; AAB58424.1; -.  
DR EMBL; U79573; AAB58425.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 263 AA; 30516 MW; E245DF7483A5A0DD CRC64;

Query Match	70.6%;	Score 142;	DB 11;	Length 263;
Best Local Similarity	64.1%;	Pred. No. 1.8e-10;		
Matches	25;	Conservative 11;	Mismatches 3;	Indels 0;
				Gaps 0.

QY 1 DSVSTSFKRLRQQLGPVTOEWFMDNEKETEGRLRQEMSKD 39  
 74 DTIGSTVSQLOERLGPLTRDFWMDNEKETDWRQEMKND 112

RESULT	4		
008855			
ID	008855	PRELIMINARY;	PRT; 263 AA.
AC	008855;		
DT	01-JUL-1997	(TREMBlrel. 04, Created)	
DT	01-JUL-1997	(TREMBlrel. 04, Last sequence update)	
DT	01-DEC-2001	(TREMBlrel. 19, Last annotation update)	
DE	APOLIPOPROTEIN A-I.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6; TISSUE=SPLEEN;		
RA	MEDLINE=98077648; PubMed=9415807;		
RT	Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;		
RT	"Repetitive elements in the third intron of murine apolipoprotein A-I		
RT	gene.";		
RL	Biochem. Mol. Biol. Int. 43:989-996(1997).		
DR	EMBL; U79574; AAB58426.1; -.		
DR	HSSP; P02647; IAVL.		
DR	InterPro; IPR000074; Apolipoprotein.		
DR	Pfam; PF01442; Apolipoprotein; 1.		
KM	Lipoprotein.		
SO	SEQUENCE 263 AA; 30544 MW; 7825DF7483A5B24A CRC64;		

Query Match	70.6%;	Score 142;	DB 11;	Length 263;
Best Local Similarity	64.1%;	Pred. No. 1.8e-10;		
Matches	25;	Conservative	11;	Mismatches 3;
				Indels 0;
				Gaps 0;

QY	1	DSVSTFSFKRLRQGLGPVTOEFWMDLEKETEGLRQEMSKD	39
		:      :   :   :   :   :   :   :   :   :   :   :   :	
Db	74	DTLGSTVSQLOERLGPLTRDFWMDLEKETDWMVRQEMNKK	112

RESULT	5	
009054		
ID	009054	PRELIMINARY;
		PRT;
		258 AA

NCBI\_TaxID=10116;  
 OX  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus  
 OC Fukayaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OS Rattus norvegicus (Rat).  
 DE APOLOPROTEIN A-1.  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 NCBI\_TaxID=10116;  
 OX  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus  
 OC Fukayaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OS Rattus norvegicus (Rat).  
 DE APOLOPROTEIN A-1.  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)

RP SEQUENCE FROM N.A.  
 RC STRAIN=WKY, AND SHRSP; TISSUE=SPLEEN;  
 RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U79578; AAB58430.1; -.  
 DR EMBL; U79577; AAB58429.1; -.  
 DR HSSP; P02647; IAVI.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 1.  
 KW Lipoprotein.  
 SQ SEQUENCE 258 AA; 29918 MW; 093E6EF2E629CDBC8 CRC64;

Query Match	66.2%;	Score 133;	DB 11;	Length 258;
Best Local Similarity	64.1%;	Pred. No. 2.6e-09;		
Matches 25; Conservative	7; Mismatches 7;	Indels 0;	Gaps 0;	

```

QY      1 DSVTSTFSKLRQLGVPVQGFWMNLNLEKETEGLRQEMSKD 39
        |::||::|:|||||::|:|||||::|:|||||::|:|||||
Db      73 DFLGSTVGRQLQLGVPVQGFWMNLNLEKETDWPFRNEMNKD 111

```

## RESULT 6





DR pfam: PF01442; Apolipoprotein; 1.  
FT NON\_TER 1  
SQ SEQUENCE 244 AA; 28157 MW; 01616ACB53F12734 CRC64;

Query Match 31.3%; Score 63; DB 4; Length 244;  
Best Local Similarity 27.8%; Pred. No. 2.7;  
Matches 10; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 3 VTSTFSKLRQGLGPVTQEFWDNLEKETEGLRQEMSK 38  
Db 122 ISASAEFLRQRLAPLAEVDVRGNLKGNTGLOKSLAE 157

RESULT 10  
ID Q67018 PRELIMINARY; PRT; 562 AA.

AC Q67018;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HEMAGGLUTININ.

GN HA.  
OS Influenza A virus (A/herring gull/DE/677/88 (H2N8)).  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group; Influenza A viruses; Influenza A virus.  
OX NCBI\_TaxID=139294;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=A/HERRING GULL/DE/677/88;  
RX MEDLINE=93276567; Pubmed=7684877;  
RA Schafer J.R., Kawaoka Y., Bean W.J., Suss J., Senne D., Webster R.G.;  
RT "Origin of the pandemic 1957 H2 influenza A virus and the persistence  
of its possible progenitors in the avian reservoir."  
RL Virology 194:781-788(1993).  
CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO  
CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).  
CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS  
(HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.

DR EMBL; L11132; AAAA3096.1; -.  
DR HSSP; P03437; IHTM.  
DR InterPro; IPR001364; Hemagglutn.  
DR InterPro; IPR000529; Ribosomal\_S6.  
DR Pfam; PF00509; Hemagglutinin; 1.  
DR PRINTS; PR00329; HEMAGGLUTN12.  
DR ProDom; PD000225; Hemagglutn; 1.  
KW Envelope protein; Glycoprotein; Hemagglutinin.  
SQ SEQUENCE 562 AA; 63104 MW; 04375F5D33DBF5A3 CRC64;

Query Match 30.1%; Score 60.5; DB 12; Length 562;  
Best Local Similarity 38.9%; Pred. No. 14;  
Matches 14; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 DSVTSTFSKLRQGLGPVTQEFWDNLEKETEGLRQEM 36  
Db 390 NKVNSVTEKMTQFGAVGKEF-SNLERLLENLKKM 424

RESULT 11  
ID Q98TG1 PRELIMINARY; PRT; 259 AA.

AC Q98TG1;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE 28KDA-1E APOLIPROTEIN.  
OS Anguilla japonica (Japanese eel).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
OC Anguillidae; Anguilla.  
OX NCBI\_TaxID=7937;

RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=21175592; Pubmed=11278178;  
RA Kondo H., Kawazoe I., Nakaya M., Kikuchi K., Aida K., Watabe S.;  
RT "The novel sequences of major plasma apolipoproteins in the eel  
Anguilla japonica."  
RL Biochim. Biophys. Acta 1531:132-142(2001).  
DR EMBL; AB046208; BAB40965.1; -.  
DR InterPro; IPR000074; Apolipoprotein.  
DR pfam; PF01442; Apolipoprotein; 1.

KW Lipoprotein.  
SQ SEQUENCE 259 AA; 29895 MW; C392ADP5BAC71FAF CRC64;

Query Match 29.6%; Score 59.5; DB 13; Length 259;  
Best Local Similarity 32.6%; Pred. No. 8.1;  
Matches 15; Conservative 7; Mismatches 13; Indels 11; Gaps 1;

QY 2 SVTSTFSKLRQGLGPVT-----QEFWDNLEKETEGLRQEM 36  
Db 170 SVEQLSHLDQDLGPSTFEELKGVESMQEFWEVVRPLDRSLQKEI 215

RESULT 12  
ID Q95K18 PRELIMINARY; PRT; 311 AA.

AC Q95K18;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 36.0 KDA PROTEIN.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;

RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TESTIS;  
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,  
RA Terao K., Sugano S.;  
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA  
libraries."  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB070009; BAB62954.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 311 AA; 36012 MW; 697B3F5A8E51ED1 CRC64;

Query Match 28.9%; Score 58; DB 6; Length 311;  
Best Local Similarity 40.5%; Pred. No. 15;  
Matches 15; Conservative 7; Mismatches 13; Indels 2; Gaps 1;

QY 2 SVTSTFSKLRQGLGPVTQEFWDNLEKETEGLRQEMSK 38  
Db 135 AMTDKFSNLRD--KKHOEMMGPIEKDNLRLRQVSK 169

RESULT 13  
ID Q9LCR0 PRELIMINARY; PRT; 90 AA.

AC Q9LCR0;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE OREFX.  
GN OREFX.  
OS Paenibacillus polymyxa (Bacillus polymyxa).  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Paenibacillus.  
OX NCBI\_TaxID=1406;



Sun Sep 22 12:10:39 2002

us-09-803-918a-2\_copy\_75\_113.rspt



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:05:16 ; Search time 139.03 Seconds  
(without alignments)  
31.158 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_75\_113

Perfect score: 201  
Sequence: 1 DSVTSTFSKLRQLGPVTQEFWDNLKETEGLRQEMSKD 39

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:\*

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	119	22	AAU30468 Novel human secret
2	201	100.0	120	22	AAU30267 Novel human secret
3	201	100.0	120	22	AAU30469 Novel human secret
4	201	100.0	166	22	AAU28372 Novel human secret
5	201	100.0	243	9	AAU81082 Sequence of mature
6	201	100.0	244	22	AAU28184 Novel human secret
7	201	100.0	264	15	AAU56863 Apo-1ipoprotein AI
8	201	100.0	264	15	AAU56864 Apo-1ipoprotein AI
9	201	100.0	267	7	AAU61079 Assumed human apol
10	201	100.0	267	9	AAU82128 Entire human prepr
11	201	100.0	267	14	AAU34032 Sequence of apo AI

12	201	100.0	267	16	AAU72705 Human apo A-I incl
13	201	100.0	267	18	AAU08602 Human apolipoprote
14	201	100.0	267	20	AAU18675 Human apolipoprote
15	201	100.0	267	22	AAU47620 Full length Apo-AI
16	201	100.0	268	9	AAU80668 Recombinant human
17	201	100.0	299	22	AAU33170 Novel human secret
18	195	97.0	221	22	AAU29835 Novel human secret
19	192	95.5	151	22	AAU02278 Human polypeptide
20	181	90.0	154	22	AAU01295 Human polypeptide
21	131.5	65.4	318	22	AAU30268 Novel human secret
22	80	39.8	16	22	AAU62607 Apolipoprotein fra
23	63.5	31.6	50	22	AAU02135 Mouse alpha helica
24	63.5	31.6	304	22	AAU02129 Mouse mature alpha
25	63.5	31.6	331	22	AAU02128 Mouse alpha helica
26	63	31.3	396	22	AAU90664 Human secreted pro
27	62	30.8	194	14	AAU39483 Human apoAIV mutei
28	62	30.8	328	14	AAU39484 Human apoAIV mutei
29	62	30.8	333	14	AAU39481 Human apoAIV mutei
30	62	30.8	333	14	AAU39488 Human apoAIV mutei
31	62	30.8	333	14	AAU39495 Human apoAIV mutei
32	62	30.8	333	14	AAU39497 Human apoAIV mutei
33	62	30.8	337	14	AAU39485 Human apoAIV mutei
34	62	30.8	342	14	AAU39487 Human apoAIV mutei
35	62	30.8	342	14	AAU39489 Human apoAIV mutei
36	62	30.8	342	14	AAU39496 Human apoAIV mutei
37	62	30.8	342	14	AAU39498 Human apoAIV mutei
38	62	30.8	363	14	AAU39478 Human apoAIV mutei
39	62	30.8	363	14	AAU39479 Human apoAIV mutei
40	62	30.8	373	14	AAU39486 Human apoAIV mutei
41	62	30.8	377	14	AAU39443 Human apolipoprote
42	62	30.8	377	14	AAU39480 Human apoAIV mutei
43	62	30.8	377	14	AAU39499 Human apoAIV mutei
44	62	30.8	377	14	AAU39500 Human apoAIV mutei
45	62	30.8	377	14	AAU39502 Human apoAIV mutei

## ALIGNMENTS

RESULT 1	AAU30468	standard; Protein; 119 AA.
ID	AAU30468	
XX	AAU30468;	
AC	18-DEC-2001	(first entry)
XX		
DT	18-DEC-2001	(first entry)
XX		
DE	Novel human secreted protein #959.	
XX		
KW	Human; vaccination; gene therapy; nutritional supplement;	
KW	stem cell proliferation; haematopoiesis; nerve tissue regeneration;	
KW	immune suppression; immune stimulation; anti-inflammatory; leukaemia.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200179449-A2.	
XX		
PD	25-OCT-2001.	
XX		
PF	16-APR-2001; 2001WO-US08656.	
XX		
PR	18-APR-2000; 2000US-0552929.	
XX		
PA	26-JAN-2001; 2001US-0770160.	
XX		
PI	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT;	
XX		
DR	WPI; 2001-611725/70.	
XX		
PT	Nucleic acids encoding a range of human polypeptides, useful in genetic	
PT	vaccination, testing and therapy -	
XX		

PS Claim 20; Page 297; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 119 AA;  
  
Query Match 100.0%; Score 201; DB 22; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.4e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DSVTSTFSKLRQLGPLYTQEFWDNLEKETEGLRQEMSKD 39  
ID AAV30267 standard; Protein; 120 AA.  
XX  
AC AAU30267;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #758.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 270; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 120 AA;  
  
Query Match 100.0%; Score 201; DB 22; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.4e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DSVTSTFSKLRQLGPLYTQEFWDNLEKETEGLRQEMSKD 39  
ID AAV30469 standard; Protein; 120 AA.  
XX  
AC AAU30469;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #960.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 297; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to

CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

XX SQ Sequence 120 AA;

Query Match 100.0%; Score 201; DB 22; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.4e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 39  
Db 52 dsvtstfsklrqlgptvgtfwdnleketeglrqemskd 90

## RESULT 4

AAU28372  
ID AAU28372 standard; Protein; 166 AA.

XX AC AAU28372;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secretory protein, Seq ID No 729.

XX KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
KW fertility; analgesic; pain; antigen.

XX OS Homo sapiens.

XX PN WO200166689-A2.

XX PD 13-SEP-2001.

XX PF 05-MAR-2001; 2001WO-US04942.

XX PR 07-MAR-2000; 2000US-0519705.

XX PR 19-MAY-2000; 2000US-0574454.

XX PR 17-JUN-2000; 2000US-0596193.

XX PR 14-JUL-2000; 2000US-0616847.

XX PR 19-SEP-2000; 2000US-0665363.

XX PR 20-OCT-2000; 2000US-0693267.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

XX PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX DR WPI; 2001-589934/66.

XX DR N-PSDB; AAS45272.

XX PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
XX PT prepared from various human tissues, for diagnosis and treatment of  
XX PT cancer, neurological, inflammatory, and autoimmune disorders -

XX PS Example 2; SEQ ID No 729; 107pp; English.

XX CC The invention relates to novel isolated human secreted polypeptides (I)  
XX CC and polynucleotides (II). (I) and (II) are useful for treating  
XX CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
XX CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
XX CC involved in increasing haematopoiesis, stem cell survival, bone growth  
XX CC and remodeling. (I), (II) and modulators of (II) are useful for

CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
CC creating transgenic animals useful for studying the in vivo activities of  
CC the polypeptide as well as for studying modulators of the polypeptides.  
CC (I) induces the proliferation of neural cells and regeneration of nerve  
CC and brain tissue and is useful for the treatment of central and  
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
CC activity, regulation of haematopoiesis and is useful for treating myeloid  
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
CC tissue growth, and in tissue repair, healing of burns, incisions,  
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative  
CC disorders, or periodontal disease. Furthermore, (I) is also useful for  
CC gut protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues, various immune deficiencies and  
CC disorders including severe combined immunodeficiency (SCID), bacterial or  
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,  
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
CC reactions and conditions, such as asthma or other respiratory problems.  
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,  
CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
CC analgesic effects or other pain reducing effects, immunoglobulin like  
CC activity and can act as an antigen in a vaccine composition to raise an  
CC immune response. AAU28020-AAU28395 represent novel human secreted protein  
XX amino acid sequences of the invention.

SQ Sequence 166 AA;

Query Match 100.0%; Score 201; DB 22; Length 166;  
Best Local Similarity 100.0%; Pred. No. 2.1e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQEFWMDNLEKETEGLRQEMSKD 39  
Db 61 dsvtstfsklrqlgptvgtfwdnleketeglrqemskd 99

## RESULT 5

AAU81082  
ID AAU81082 standard; Protein; 243 AA.

XX AC AAU81082;

XX DT 14-JAN-1991 (first entry)

XX DE Sequence of mature human apolipoprotein AI (apoAI).

XX KW Atherosclerosis; therapy; cardiovascular disease.

XX OS Homo sapiens.

XX PN WO8803166-A.

XX PD 05-MAY-1988.

XX PF 21-OCT-1987; 87WO-EP00621.

XX PR 23-OCT-1986; 86GB-0025435.

XX PA (FARM ) FARMITALIA C ERBA SPA.

XX PI Lorenzetti R, Monaco L, Soria M, Palomba R, Isacchi A, Sarmientos P;

XX DR WPI; 1988-133240/19.

XX DR N-PSDB; AAN80243.

XX PT Recombinant human apolipoprotein AI -  
XX PT used to lower plasma cholesterol and/or tri glyceride levels and  
XX PT to combat atherosclerosis and cardiovascular diseases

PS Disclosure; Fig 1; 51pp; English.

XX

CC The protein comprising apo AI genetic variants may be used to lower

CC plasma cholesterol and/or triglyceride levels. They may also be used to

CC combat atherosclerosis and cardiovascular diseases such as coronary

CC heart disease. Prefd. proteins are Met-apo AI, Met-apo AI-T6, Met-apo

CC AI-MI and Met-apo AI-T6/MI.

XX

SQ Sequence 243 AA;

Query Match 100.0%; Score 201; DB 9; Length 243;

Best Local Similarity 100.0%; Pred. No. 3.3e-18;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPTQGFWDNLEKETEGLRQEMSKD 39

Db 51 dsvtstfskrlreglgyptqgfwdnleketeglrqemskd 89

RESULT 6

AAU28184

ID AAU28184 standard; Protein; 244 AA.

XX

AC AAU28184;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secretory protein, Seq ID No 353.

XX

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;

KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;

KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;

KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;

KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;

KW gut protection; lung; liver fibrosis; immune deficiency; infection;

KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;

KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;

KW fertility; analgesic; pain; antigen.

XX

OS Homo sapiens.

XX

PN WO200166689-A2.

XX

PD 13-SEP-2001.

XX

PF 05-MAR-2001; 2001WO-US04942.

XX

PR 07-MAR-2000; 2000US-0519705.

PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0616847.

PR 19-SEP-2000; 2000US-0665363.

PR 20-OCT-2000; 2000US-0693267.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX

DR WPI; 2001-589934/66.

DR N-PSDB; AAS45084.

XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries

PT prepared from various human tissues, for diagnosis and treatment of

PT cancer, neurological, inflammatory, and autoimmune disorders -

XX

PS Example 5; SEQ ID No 353; 107pp; English.

XX

CC The invention relates to novel isolated human secreted polypeptides (I)

CC and polynucleotides (II). (I) and (II) are useful for treating

CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

CC ischaemia, reperfusion injury, shock, sepsis, immune responses, and is

CC involved in increasing haematopoiesis, stem-cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for

CC prophylaxis or treatment of one or more cancers. (II) is also useful for

CC creating transgenic animals useful for studying the in vivo activities of

CC the polypeptide as well as for studying modulators of the polypeptides.

CC (I) induces the proliferation of neural cells and regeneration of nerve

CC and brain tissue and is useful for the treatment of central and

CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,

CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic

CC activity, regulation of haematopoiesis and is useful for treating myeloid

CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia

CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve

CC tissue growth, and in tissue repair, healing of burns, incisions,

CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative

CC disorders, or periodontal disease. Furthermore, (I) is also useful for

CC gut protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues, various immune deficiencies and

CC disorders including severe combined immunodeficiency (SCID), bacterial or

CC fungal infections, autoimmune disorders e.g. multiple sclerosis,

CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic

CC reactions and conditions, such as asthma or other respiratory problems.

CC In addition, (I) affects biorhythms or circadian cycles of rhythms,

CC fertility, metabolism, catabolism, anabolism, storage or elimination of

CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides

CC analgesic effects or other pain reducing effects, immunoglobulin like

CC activity and can act as an antigen in a vaccine composition to raise an

CC immune response. AAU28020-AAU28395 represent novel human secreted protein

XX

amino acid sequences of the invention.

SQ Sequence 244 AA;

Query Match 100.0%; Score 201; DB 22; Length 244;

Best Local Similarity 100.0%; Pred. No. 3.3e-18;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPTQGFWDNLEKETEGLRQEMSKD 39

Db 52 dsvtstfskrlreglgyptqgfwdnleketeglrqemskd 90

RESULT 7

AAR56863

ID AAR56863 standard; Protein; 264 AA.

XX

AC AAR56863;

XX

DT 26-JAN-1995 (first entry)

XX

DE Apo-lipoprotein AI-M.

XX

KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid pKP683;

KW vector; atherosclerosis; cardiovascular disease.

XX

OS Homo sapiens.

XX

PN WO9413819-A.

XX

PD 23-JUN-1994.

XX

PF 09-DEC-1993; 93WO-SE01061.

XX

PR 11-DEC-1992; 92SE-0003753.

XX

PA (KABI ) KABI PHARMACIA AB.

XX

PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;

PI Sejlitz T;

XX

DR WPI; 1994-217892/26.

DR N-PSDB; AAQ68357.

XX



PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 3; 33pp; English.

CC Plasmid PKP683 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of PKP683 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.

SQ Sequence 264 AA;

Query Match 100.0%; Score 201; DB 15; Length 264;  
Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQGFWDNLEKETEGLRQEMSKD 39  
Db 72 dsvtstfkslreqlqpvgtqfwdnleketeglrqemskd 110

## RESULT 8

AAR56864 AAR56864 standard; Protein; 264 AA.

AC AAR56864;

DT 26-JAN-1995 (first entry)

DE Apo-lipoprotein AI-M.

KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid PKP764;  
KW vector; atherosclerosis; cardiovascular disease.

OS Homo sapiens.

PN W09413819-A.

PD 23-JUN-1994.

PF 09-DEC-1993; 93WO-SE01061.

PR 11-DEC-1992; 92SE-0003753.

PA (KABI ) KABI PHARMACIA AB.

PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
PI Sejlitz T;

DR WPI; 1994-217892/26.  
DR N-PSDB; AAQ68358.

PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 4; 33pp; English.

CC Plasmid PKP764 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of PKP764 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.

SQ Sequence 264 AA;

Query Match 100.0%; Score 201; DB 15; Length 264;  
Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQGFWDNLEKETEGLRQEMSKD 39  
Db 72 dsvtstfkslreqlqpvgtqfwdnleketeglrqemskd 110

## RESULT 9

AAP61079 AAP61079 standard; Protein; 267 AA.

AC AAP61079;

DT 07-OCT-1991 (first entry)

DE Assumed human apolipoprotein A-1 derivative gene product.

KW Hyperlipaemia; arteriosclerosis.

OS Homo sapiens.

PN JP61096998-A.

PD 15-MAY-1986.

PF 16-OCT-1984; 84JP-0216988.

PR 16-OCT-1984; 84JP-0216988.

PA (MITU ) MITSUBISHI CHEM IND KK.

DR WPI; 1986-165025/26.

DR N-PSDB; AAN60886.

PT Human apo-lipoprotein A-1 (deriv.) prepn. - by providing DNA  
PT fragment in cloning site downstream of expression vector promoter  
PT and introducing into host microorganism.

PS Disclosure; Fig 2; 9pp; Japanese.

CC The human apolipoprotein may be produced by a suitable transformed  
CC host, it is effective in treating hyperlipaemia and arteriosclerosis.

SQ Sequence 267 AA;

Query Match 100.0%; Score 201; DB 7; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPVTQGFWDNLEKETEGLRQEMSKD 39  
Db 75 dsvtstfkslreqlqpvgtqfwdnleketeglrqemskd 113

## RESULT 10

AAP82128 AAP82128 standard; protein; 267 AA.

AC AAP82128;

DT 24-OCT-1990 (first entry)

DE Entire human preproapoprotein AI.

KW human preproapoprotein AI; high density lipoprotein deficiency; ss.

OS synthetic.

FT Key Location/Qualifiers

FT Peptide 1..18

FT Peptide /label=precursor

FT Peptide 19..24

FT /label=propeptide

FT	Protein	25..267
FT	/label=mature apoprotein	
PN	EP293357-A.	
XX		
PD	30-NOV-1988.	
XX		
PF	24-MAY-1988;	88EP-0870095.
XX		
PR	28-MAY-1987;	87GB-0012540.
XX		
PA	(UNIO ) UCB SA.	
XX		
PI	Bollen A, Gobert J, Wulfert E;	
XX		
DR	WPI; 1988-339891/48.	
DR	N-PSDB; AAN82064.	
XX		
PT	New DNA encoding human preproapoprotein A1 -	
PT	modified to eliminate hairpin structures	
XX		
PS	Disclosure; ; p; French.	
XX		
CC	The cDNA 878bp fragment encoding preproapoprotein A1 was detected	
CC	in clone pULB1609 derived from human liver cells.	
CC	See also AAN81258.	
XX		
SQ	Sequence	267 AA;
	Query Match	100.0%; Score 201; DB 9; Length 267;
	Best Local Similarity	100.0%; Pred. No. 3.6e-18;
	Matches	39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DSVSTFSKLRQQLGPVTQEFWDLNLEKETEGLRQEMSKD	39
Db	75 dsvstfskrlreqqlgptvqefwlnleketeglrqemskd	113
	RESULT 11	
	AAR34032	
ID	AAR34032 standard; Protein; 267 AA.	
XX		
AC	AAR34032;	
XX		
DT	13-AUG-1993 (first entry)	
XX		
DE	Sequence of apo A1.	
XX		
KW	Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9307165-A.	
XX		
PD	15-APR-1993.	
XX		
PE	09-OCT-1992;	92WO-US08634.
XX		
PR	09-OCT-1991;	91US-0774633.
PR	08-OCT-1992;	92US-0555555.
PR	28-JUN-1992;	92US-0901706.
XX		
PA	(SCRI ) SCRIPPS RES INST.	
XX		
PI	Curtiss LK, Koduri KR, Smith RS, Wiltztlum JL, Young SG;	
XX		
DR	WPI; 1993-134378/16.	
DR	N-PSDB; AAQ40030.	
XX		
PT	Polyptide mimic of native apo B-100 and native apo A-I - useful	
PT	in assays for LDL and HDL in plasma samples	
XX		

PS	Claim 19; Pages 105-106; 137pp; English.
XX	
CC	The inventors claim a portion of the polypeptide contg. apo B-100
CC	that immunoreacts with antibodies secreted by the hybridoma MB47
CC	having ATCC Accession No. 8746. Polypeptides specifically claimed
CC	include residues 217-297, 216-310, 216-331, 216-352, 216-377, 1-377,
CC	205-297, 173-297, 140-297. DNA sequences encoding the polypeptides
CC	are also claimed. Also claimed are a fusion polypeptide that
CC	contains: (a) a first amino acid residue sequence up to 250 residues
CC	in length that includes residues 120-135 of apo A-I, (b) a second
CC	amino acid residue sequence up to 375 residues in length that
CC	includes residues 217-297 of apo B-100 and DNA encoding it.
XX	
XX	Sequence 267 AA;
XX	
Query Match	100.0%; Score 201; DB 14; Length 267;
Best Local Similarity	100.0%; Pred. No. 3.6e-18;
Matches 39, Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 DSVTSTFSKLRQQLGPVTQEEFWDNLEKETEGLRQEMSKD 39
Db	75 dsvststfskrlreglqgvtqefwfnleketeglrqemskd 113
RESULT 12	
ID	AAR72705 standard; Protein; 267 AA.
XX	
AC	AAR72705;
XX	
DT	31-OCT-1995 (first entry)
XX	
DE	Human apo A-I including signal and propeptide sequences.
XX	
KW	Apo A-I; LDL cholesterol; low density lipoprotein; lipid.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	Peptide 1..18
FT	/label= presignal
FT	Peptide 19..24
FT	/label= propeptide
FT	Peptide 120..135
FT	/label= claimed
FT	/note= "as part of fusion polypeptide"
FT	Peptide 19..240
FT	/label= claimed
FT	/note= "as part of fusion polypeptide"
XX	
PN	US5408038-A.
XX	
PD	18-APR-1995.
XX	
PF	09-OCT-1991; 91US-0774633.
XX	
PR	09-OCT-1991; 91US-0774633.
PR	18-JUN-1992; 92US-0901706.
PR	08-OCT-1992; 92US-0959946.
XX	
PA	(SCRI ) SCRIPPS RES INST.
XX	
PI	Curtiss LK, Koduri KR, Smith RS, Witzum JL, Young SG;
XX	
DR	WPI; 1995-161146/21.
DR	N-PSDB; AAQ89634.
XX	
PT	New apo:lipoprotein B-100 peptide(s) and fusion peptide(s) - used
PT	in assay systems for detecting LDL and HDL cholesterol levels in
PT	body fluids.
XX	
PS	Claim 10; Fig 2; 41pp; English.

XX AAA89634 and AAR72705 depict the AA sequence of human apo A-I and its  
CC corresp. cDNA, including presignal residues and propeptide  
CC residues, according to Seilhamer et al.. DNA 3(4):309 (1984).  
CC A dispersible apo A-I/B-100 fusion polypeptide is claimed which  
CC contains a first AA sequence of apo A-I and that includes at  
CC least AA sequence positions 120-135 (see AAR72606) and which reacts  
CC with pan anti-apo A-I antibodies such as: AI-4 ATCC HB8744; AI-7  
CC ATCC HB 8745; AI-9 ATCC HB 8741; AI-10 ATCC HB 9200; AI-11 ATCC  
CC HB 9201; AI-12 ATCC HB 9202; AI-13 ATCC HB 9203; AI-14 ATCC HB  
CC 9204; AI-18 ATCC HB 9507.  
XX  
SQ Sequence 267 AA;

Query Match 100.0%; Score 201; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
Db 75 dsvtstfkslreqlgptqefwdnleketeglrqemskd 113

RESULT 13  
AAW08602  
ID AAW08602 standard; Protein; 267 AA.

XX AAW08602;

DT 04-SEP-1997 (first entry)

DE Human apolipoprotein A-1 variant "Paris" protein sequence.

KW Human; apolipoprotein A-1; variant; mutation; serum lipid; cholesterol;  
high density lipoprotein; triglyceride; symptom; cardiovascular disease;  
atherosclerosis; restenosis; myocardial infarct; angina; death; stroke;  
cardiac decompensation; metabolic deficit.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Peptide 1..24  
FT /note= "prepro peptide"  
FT Region 172..178

FT /note= "this sequence which contains the mutated amino  
acid residue is claimed - claim 2"

FT Misc-difference 175  
FT /note= "changed from Arg residue in wild type protein,  
due to a C to T transition mutation"

PN WO9637608-A1.

PD 28-NOV-1996.

PF 20-MAY-1996; 96WO-FR00747.

PR 22-MAY-1995; 95FR-0006061.

PA (INSP ) INST PASTEUR LILLE.  
PA (RHON ) RHONE-POULENC RORER SA.  
PA (UYPA-) UNIV CURIE PARIS VI P & M.

PI Benoit P, Bruckert E, Deneffe P, Duverger N, Fruchart J;  
PI Luc G, Turping, Assmann G, Funke H;

DR WPI; 1997-021218/02.  
DR N-PSDB; AAT43691.

PT New variant of human apoA-1 with Cys at position 151 - has  
PT anti-atherogenic activity for treatment and prevention of  
PT cardiovascular disease

PS Claim 2; Page -: 58pp; French.

XX This is the amino acid sequence of a human apolipoprotein A-1 variant  
CC designated the "Paris" variant which has a Cys replacing the Arg residue  
CC at position 151. The substitution is generated by a mutation of  
CC the C nucleotide at position 523 in the wild type gene to a T residue,  
CC changing the encoded residue from an Arg to a Cys. The gene was isolated  
CC from a patient with an unusual pattern of serum lipids i.e. low levels of  
CC apoA-1 and high density lipoprotein (HDL)-cholesterol and high  
CC triglycerides, but showing symptoms of atherosclerosis. The new variant  
CC protein has anti-atherogenic activity so is useful for treatment and  
CC prevention of cardiovascular diseases such as atherosclerosis,  
CC restenosis, myocardial infarct, angina, sudden death, stroke and cardiac  
CC decompensation, or more generally any condition involving genetic or  
CC metabolic deficit of apoA-1.  
CC Note: this sequence is not given in the specification but is generated  
CC from the wild type apoA-1 gene disclosed in the specification and has  
CC the appropriate amino acid changed.

SQ Sequence 267 AA;

Query Match 100.0%; Score 201; DB 18; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
Db 75 dsvtstfkslreqlgptqefwdnleketeglrqemskd 113

RESULT 14

AAV18675  
ID AAV18675 standard; Protein; 267 AA.

AC AAV18675;

DT 09-JUL-1999 (first entry)

DE Human apolipoprotein AI protein sequence.

KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.

OS Homo sapiens.

PN WO9916409-A2.

PD 08-APR-1999.

PF 28-SEP-1998; 98WO-US20329.

PR 29-SEP-1997; 97US-0940136.

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (DUFO/) DUFOURCO J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;  
PI Sekul R;

DR WPI; 1999-254921/21.  
DR N-PSDB; AAX55971.

PT Nucleic acid encoding apolipoprotein A-I agonist peptides  
PT Example; Fig 1; 232pp; English.  
PT The present invention describes a nucleic acid (A) encoding an

```
CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,
CC which forms an amphipathic alpha-helix in presence of lipids. (A)',
CC optionally as a complex with lipids, and host cells that contain (A)',
CC are useful for gene therapy, or prevention, of diseases associated with
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I
CC deficiency, hypertriglyceridemia and metabolic syndrome, also to treat
CC endotoxemia (septic shock). Host cells containing (A) can also be used
CC to study the role of apoA-I in lipid metabolism. (B) can be used
CC diagnostically, e.g. to measure serum HDL (particularly its
CC subpopulation involved in retrograde cholesterol transport) and for
CC imaging the circulatory system or HDL accumulations at fatty streaks.
CC The present sequence represents human apoA-I.
XX
SQ Sequence 267 AA;

Query Match 100.0%; Score 201; DB 20; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.6e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPTQEFWDNLEKEFTGLRQEMSKD 39
   |||||
Db 75 dsvtstfslkrlqglpvtqefwdnleketeglrqemskd 113

RESULT 15
AAB47620
ID AAB47620 standard; Protein; 267 AA.
XX
XX AAB47620;
AC
XX
DT 21-JAN-2002 (first entry)
XX
DE Full length Apo-A1.
XX
KW Apolipoprotein; Apo-A1; Apo-A-I fragment T-cell activation inhibitor;
KW AFT1; monocyte; IL-1; interleukin 1; TNF; tumour necrosis factor;
KW acute pancreatitis; Alzheimer's disease; asthma; cancer; fever;
KW inflammatory bowel disease; ischemia; multiple sclerosis; osteoporosis;
KW Parkinson's disease; psoriasis; probe.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Binding-site 44..65
FT Binding-site /label= Helical lipid binding domain
FT Binding-site 220..241
FT Binding-site /label= Helical lipid binding domain
FT Domain 74..111
FT /note= "Involved in lipoprotein-mediated cholesterol
FT efflux from monocytes"
FT Binding-site 149..219
FT /label= Receptor binding domain
FT Domain 99..120
FT /label= Major antigenic epitope domain
FT Domain 99..143
FT /label= Hinged domain
FT Domain 66..120
FT /label= Phylogenetically conserved domain
FT Domain 90..111
FT /note= "Involved in lectin-cholesterol acyltransferase
FT activity"
FT Domain 44..65
FT /label= Amphipathic helix
FT Domain 66..98
FT /label= Amphipathic helix
FT Domain 99..120
FT /label= Amphipathic helix
FT Domain 121..142
FT /label= Amphipathic helix
FT Domain 143..164
FT /label= Amphipathic helix
FT Domain
```

```
FT Domain 165..208
FT /label= Amphipathic helix
FT Domain 209..219
FT /label= Amphipathic helix
FT Domain 220..241
FT /label= Amphipathic helix
FT Peptide 25..194
FT /label= AFTI
FT /note= "18 kD N-terminal fragment"
FT Peptide 25..144
FT /label= AFTI
FT /note= "13 kD N-terminal fragment"
FT Peptide 156..267
FT /label= AFTI
FT /note= "13 kD C-terminal fragment"
XX
PN WO200168852-A2.
XX
XX 20-SEP-2001.
PD
XX 13-MAR-2001; 2001WO-US07826.
PF 13-MAR-2000; 2000US-189008P.
PR 13-MAR-2000; 2000US-189008P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Edwards CK, Burger D, Dayer J, Kohno T;
XX WPI; 2001-596908/67.
DR N-PSDB; AAH43623.
XX
PT Apo-A-I fragment T-cell activation inhibitor (AFT1) polynucleotides,
PT useful for treating, diagnosing, ameliorating diseases associated with
PT IL-1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease
PT and asthma -
XX
PS Claim 1; Fig 1A; 132pp; English.
XX
XX This sequence shows full length apolipoprotein (Apo-A1). Fragments
XX of Apo-A1 may be used as Apo-A-I fragment T-cell activation inhibitors
XX (AFT1). These fragments are selected from an 18 kD N-terminal fragment
XX (amino acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144)
XX and a 13 kD C-terminal fragment (amino acids 156-267). The AFTI
XX polypeptides and polynucleotides are useful for regulating T-cell
XX mediated activation of monocytes and for treating, diagnosing,
XX ameliorating diseases associated with IL-1 and/or TNF activity.
XX The diseases are acute pancreatitis, Alzheimer's disease, asthma,
XX cancer, fever, inflammatory bowel disease, ischemia, multiple
XX sclerosis, osteoporosis, Parkinson's disease or psoriasis. Numerous
XX examples of other diseases are given in the specification.
XX The AFTI nucleic acids are useful as hybridization probes in diagnostic
XX assays to test for the presence of an AFTI or Apo-A-I DNA in mammalian
XX tissue or bodily fluid samples.
XX
SQ Sequence 267 AA;

Query Match 100.0%; Score 201; DB 22; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.6e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

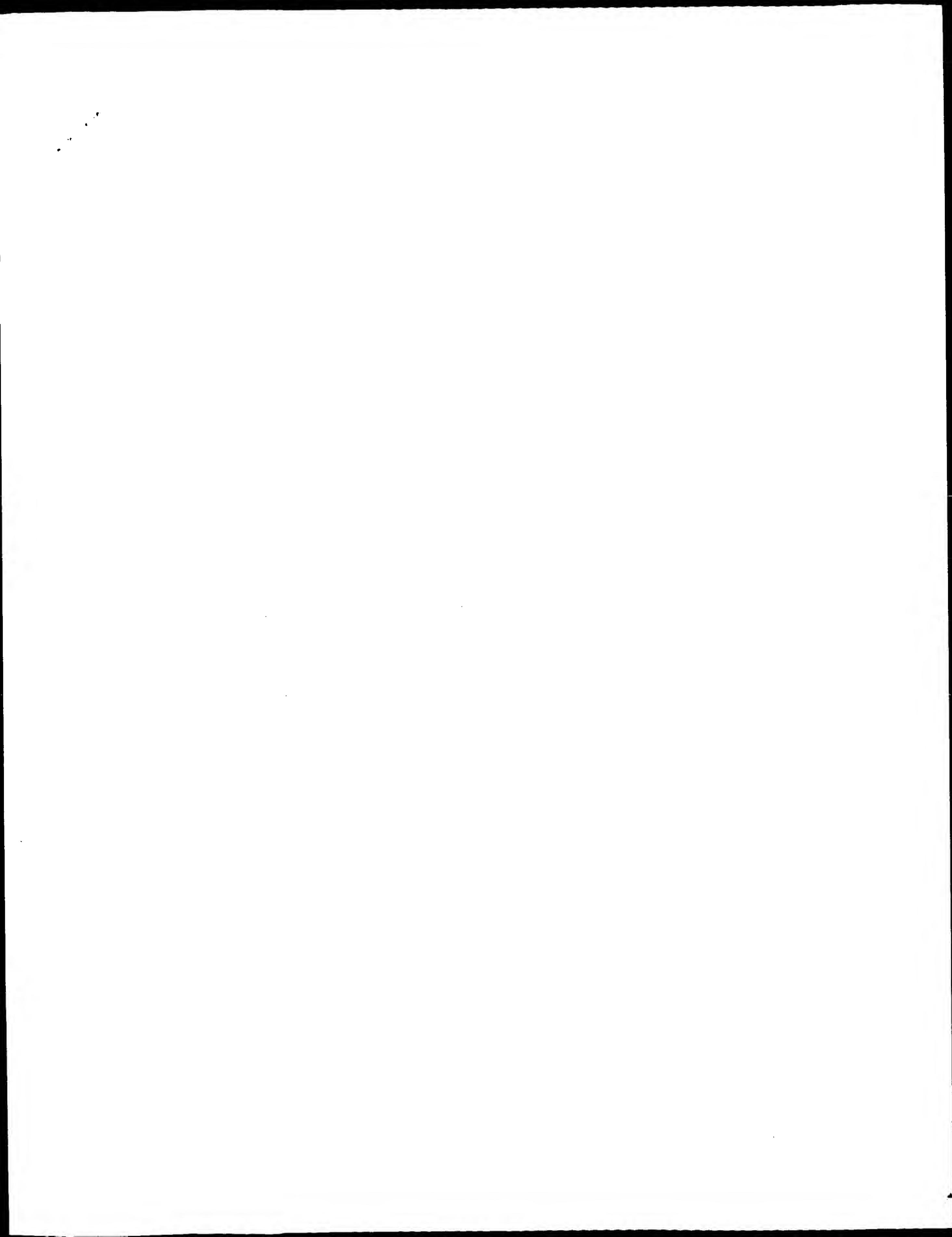
QY 1 DSVTSTFSKLRQLGPTQEFWDNLEKEFTGLRQEMSKD 39
   |||||
Db 75 dsvtstfslkrlqglpvtqefwdnleketeglrqemskd 113
```

Search completed: September 22, 2002, 12:05:16  
Job time: 291 sec



Sun Sep 22 12:10:34 2002

us-09-803-918a-2\_copy\_75\_113.rag



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:02:53 ; Search time 51.49 Seconds

(without alignments)  
18.501 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_75\_113

Perfect score: 201  
Sequence: 1 DSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKD 39

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	200	4	US-08-952-796-15
2	201	100.0	264	1	US-08-448-606-6
3	201	100.0	267	1	US-07-959-946-3
4	201	100.0	267	1	US-08-333-577-3
5	201	100.0	267	4	US-08-952-796-2
6	201	100.0	267	5	PCT-US92-08634-3
7	59	29.4	317	4	US-08-949-155-6
8	56	27.9	22	3	US-08-940-095-73
9	56	27.9	22	3	US-08-940-093-73
10	56	27.9	22	3	US-08-940-096-73
11	56	27.9	22	4	US-09-465-719-73
12	56	27.9	22	4	US-09-453-605-73
13	54.5	27.1	348	1	US-08-229-781-50
14	54.5	27.1	348	1	US-08-630-918-50
15	54.5	27.1	348	4	US-09-004-422-50
16	54.5	27.1	474	3	US-09-022-699-2
17	51.5	25.6	593	1	US-08-202-389-12
18	51.5	25.6	593	1	US-08-018-129-5
19	51.5	25.6	593	2	US-08-448-250-5
20	51.5	25.6	593	2	US-07-928-930A-8
21	51	25.4	22	1	US-08-288-568-8
22	51	25.4	22	1	US-08-487-461-8
23	51	25.4	22	1	US-08-432-691-8
24	51	25.4	22	1	US-08-487-459-8
25	51	25.4	110	1	US-07-849-389-7
26	51	25.4	220	2	US-08-726-306A-29
27	51	25.4	317	1	US-07-709-949-2

28	50.5	25.1	713	1	US-08-190-802A-63	Sequence 63, Appl
29	50.5	25.1	713	4	US-08-477-346-63	Sequence 63, Appl
30	50.5	25.1	713	4	US-08-473-089-63	Sequence 6, Appl
31	49.5	24.6	366	1	US-08-004-492-8	Sequence 8, Appl
32	49	24.4	191	4	US-08-842-306B-6	Sequence 6, Appl
33	49	24.4	191	4	US-08-838-973B-6	Sequence 11, Appl
34	49	24.4	599	2	US-08-846-526-11	Sequence 2, Appl
35	49	24.4	599	4	US-08-172-339-2	Sequence 75, Appl
36	48	23.9	22	3	US-08-940-095-75	Sequence 75, Appl
37	48	23.9	22	3	US-08-940-093-75	Sequence 75, Appl
38	48	23.9	22	3	US-08-940-096-75	Sequence 75, Appl
39	48	23.9	22	4	US-09-465-719-75	Sequence 75, Appl
40	48	23.9	22	4	US-09-453-605-75	Sequence 4, Appl
41	48	23.9	120	2	US-08-917-456-4	Sequence 4, Appl
42	48	23.9	120	4	US-09-229-804-4	Sequence 2, Appl
43	48	23.9	158	2	US-08-917-456-2	Sequence 2, Appl
44	48	23.9	158	4	US-09-229-804-2	Sequence 5, Appl
45	48	23.9	203	1	US-08-216-593-5	

## ALIGNMENTS

RESULT 1  
US-08-952-796-15  
; Sequence 15, Application US/08952796  
; Patent No: 6258596  
; GENERAL INFORMATION:  
; APPLICANT: BENOIT, Patrick  
; APPLICANT: BRUCKERT, Eric  
; APPLICANT: DENEFFE, Patrice  
; APPLICANT: DUBERGER, Nicolas  
; APPLICANT: FRUCHART, Jean-Charles  
; APPLICANT: LUC, Gerald  
; APPLICANT: TURPIN, Gerard  
; APPLICANT: ASSMANN, Gerd  
; APPLICANT: FUNKE, Harald  
; TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPROTEIN A-I  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, Mailstop 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/952,796  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/06061  
; FILING DATE: 22-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO FR96/00747  
; FILING DATE: 20-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fehner Esq., Paul F.  
; REGISTRATION NUMBER: 35,135  
; REFERENCE/DOCKET NUMBER: ST95031-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 200 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
US-08-952-796-15

Query Match 100.0%; Score 201; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 1.4e-19;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
DB 8 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 46

## RESULT 2

US-08-448-606-6  
; Sequence 6, Application US/08448606  
; Patent No. 572114  
; GENERAL INFORMATION:  
; APPLICANT: Abrahams n, Lars  
; APPLICANT: Holmgren, Erik  
; APPLICANT: Kalder n, Christina  
; APPLICANT: Lake, Mats  
; APPLICANT: Mikaelsson, sa  
; APPLICANT: Sejlitz, Torsten  
; TITLE OF INVENTION: Expression System For Producing  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pollock, Vande Sande & Priddy  
; STREET: 1990 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448, 606  
; FILING DATE: 25-AUG-1995  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/SE93/01061  
; FILING DATE: 09-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9203753-0  
; FILING DATE: 11-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amernick, Burton A.  
; REGISTRATION NUMBER: 24,852  
; REFERENCE/DOCKET NUMBER: 0151/00121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)331-7111  
; TELEFAX: (202) 293-6229  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-448-606-6

Query Match 100.0%; Score 201; DB 1; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2e-19;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
DB 72 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 110

RESULT 3  
US-07-959-946-3  
; Sequence 3, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959, 946  
; FILING DATE: 19921008  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-946-3

Query Match 100.0%; Score 201; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2e-19;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
DB 75 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 113

RESULT 4  
US-08-333-577-3  
; Sequence 3, Application US/08333577  
; Patent No. 5786206  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:

OY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
DB 72 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 110



ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 No. 5786206th Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,577  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: SCRF 234.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-333-577-3

Query Match 100.0%; Score 201; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2e-19;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
Db 75 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 113

RESULT 5  
US-08-952-796-2  
Sequence 2, Application US/08952796  
Patent No. 6258596  
GENERAL INFORMATION:  
APPLICANT: BENOIT, Patrick  
APPLICANT: BRUCKERT, Eric  
APPLICANT: DENEFFLE, Patrice  
APPLICANT: DUBERGER, Nicolas  
APPLICANT: FRUCHART, Jean-Charles  
APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Behlner Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-2

Query Match 100.0%; Score 201; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2e-19;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 39  
|||||  
Db 75 DSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKD 113

RESULT 6  
PCT-US92-08634-3  
Sequence 3, Application PC/TUS9208634  
GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Witzum, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 North Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08634  
FILING DATE: 19921009  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids

## RESULT 8

APPLICANT: Dasseux, Jean-Louis  
APPLICANT: Sekul, Renate  
APPLICANT: Buttner, Klaus  
APPLICANT: Cornut, Isabelle  
APPLICANT: Metz, Gunther

;; TITLE OF INVENTION: APOLIPROTEIN A-I AGONISTS  
;; NUMBER OF INVENTIONS: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
;; NUMBER OF SEQUENCES: 258  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds LLP  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10036-2811  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq Version 2.0  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/940,093  
;; FILING DATE: 29-SEP-1997  
;; CLASSIFICATION: 435  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 009196-0006-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-493-4935  
;; TELEFAX: 650-493-5556  
;; TELEX: 66141 PENNIE  
;;  
;; INFORMATION FOR SEQ ID NO: 73:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 22 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: No. 6037323e  
;; FEATURE:  
;; NAME/KEY: Other  
;; LOCATION: 13  
;; OTHER INFORMATION: Xaa = Alb  
;;  
US-08-940-093-73

Query Match 27.9%; Score 56; DB 3; Length 22;  
Best Local Similarity 42.9%; Pred. No. 0.26;  
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36  
Db 1 PVLDEFWEKLNEXLEALKQKL 21

RESULT 10  
US-08-940-096-73  
; Sequence 73, Application US/08940096  
; Patent No. 6046166  
; GENERAL INFORMATION:  
; APPLICANT: Dasseux, Jean-Louis  
; APPLICANT: Sekul, Renate  
; APPLICANT: Butner, Klaus  
; APPLICANT: Cornut, Isabelle  
; APPLICANT: Metz, Gunther  
; TITLE OF INVENTION: APOLIPROTEIN A-I AGONISTS  
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
; NUMBER OF SEQUENCES: 258  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq Version 2.0  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/940,096  
;; FILING DATE: 29-SEP-1997  
;; CLASSIFICATION: 530  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 009196-0005-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-493-4935  
;; TELEFAX: 650-493-5556  
;; TELEX: 66141 PENNIE  
;;  
;; INFORMATION FOR SEQ ID NO: 73:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 22 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: No. 6046166e  
;; FEATURE:  
;; NAME/KEY: Other  
;; LOCATION: 13  
;; OTHER INFORMATION: Xaa = Alb  
;;  
US-08-940-096-73

Query Match 27.9%; Score 56; DB 3; Length 22;  
Best Local Similarity 42.9%; Pred. No. 0.26;  
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36  
Db 1 PVLDEFWEKLNEXLEALKQKL 21

RESULT 11  
US-09-465-719-73  
; Sequence 73, Application US/09465719  
; Patent No. 6265377  
; GENERAL INFORMATION:  
; APPLICANT: Dasseux, Jean-Louis  
; APPLICANT: Sekul, Renate  
; APPLICANT: Butner, Klaus  
; APPLICANT: Cornut, Isabelle  
; APPLICANT: Metz, Gunther  
; TITLE OF INVENTION: APOLIPROTEIN A-I AGONISTS  
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
; NUMBER OF SEQUENCES: 258  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/465,719  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/940,093  
FILING DATE: 29-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 009196-0006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 73:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6265377e  
FEATURE:  
NAME/KEY: Other  
LOCATION: 13  
OTHER INFORMATION: Xaa = Alb  
US-09-465-719-73

Query Match 27.9%; Score 56; DB 4; Length 22;  
Best Local Similarity 42.9%; Pred. No. 0.26;  
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36  
|||:|:|:|:|:  
Db 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 12  
US-09-453-605-73  
Sequence 73, Application US/09453605  
Patent No. 6329341  
GENERAL INFORMATION:  
APPLICANT: Dasseux, Jean-Louis  
Sekul, Renate  
Bultner, Klaus  
Cornut, Isabelle  
Metz, Gunther  
Dufourcq, Jean  
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS  
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
NUMBER OF SEQUENCES: 258  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453,605  
FILING DATE: 26-No. 6329341-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/940,095  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 009196-0004-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556

TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 73:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6329341e  
FEATURE:  
NAME/KEY: Other  
LOCATION: 13  
OTHER INFORMATION: Xaa = Alb  
SEQUENCE DESCRIPTION: SEQ ID NO: 73:  
US-09-453-605-73

Query Match 27.9%; Score 56; DB 4; Length 22;  
Best Local Similarity 42.9%; Pred. No. 0.26;  
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36  
|||:|:|:|:|:  
Db 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 13  
US-08-229-781-50  
Sequence 50, Application US/08229781  
Patent No. 5589174  
GENERAL INFORMATION:  
APPLICANT: Yoshinobu OKUNO et al.  
TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/229,781  
FILING DATE: April 19, 1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/054,016  
FILING DATE: April 29, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:



INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-08-229-781-50

Query Match 27.1%; Score 54.5; DB 1; Length 348;  
Best Local Similarity 38.9%; Pred. No. 9.4;  
Matches 14; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

OY 1 DSVTSTFSKRLRQLGPTVQEFWMDNLEKETEGLRQEM 36  
DB 176 NKVNSVIEKINTQFEAVGKEF-GNLEKRLNENLNKKM 210

RESULT 14

US-08-630-918-50  
Sequence 50, Application US/08630918  
Patent No. 5631350

GENERAL INFORMATION:

APPLICANT: Yoshinobu OKUNO et al.

TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,918

FILING DATE: April 5, 1996

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/229,781

FILING DATE: April 19, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/054,016

FILING DATE: April 29, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-630-918-50

Query Match 27.1%; Score 54.5; DB 1; Length 348;  
Best Local Similarity 38.9%; Pred. No. 9.4;  
Matches 14; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

OY 1 DSVTSTFSKRLRQLGPTVQEFWMDNLEKETEGLRQEM 36  
DB 176 NKVNSVIEKINTQFEAVGKEF-GNLEKRLNENLNKKM 210

RESULT 15

US-09-004-422-50  
Sequence 50, Application US/09004422  
Patent No. 6337070

GENERAL INFORMATION:

APPLICANT: Yoshinobu OKUNO et al.

TITLE OF INVENTION: POLYPEPTIDES FOR USE IN GENERATING

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

STREET: 2033 K Street, N.W., #800

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/004,422

FILING DATE: January 8, 1998

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/443,862

FILING DATE: May 22, 1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/229,781

FILING DATE: April 19, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-721-8200

TELEFAX: 202-721-8250

TELEX:

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-004-422-50

```

```

Query Match          27.1%; Score 54.5; DB 4; Length 348;
Best Local Similarity 38.9%; Pred. No. 9.4;
Matches 14; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 1 DSVTSTFSKLRQGLGPVTQEFMDNLEKETEGLRQEM 36
   :| | | | | | | | | | | | | | | | | |
Db 176 NKVNSVIEKINTQFEAVGKEF-GNLEKRLNENLNKKM 210

```

Search completed: September 22, 2002, 12:02:54  
Job time: 254 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:06:29 : Search time 66.91 Seconds  
(without alignments)  
160.843 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_156\_267  
Perfect score: 556  
Sequence: 1 OKLHELQEKLSPLGEMRDR.....SFKVSLALAEYTKKLTNQ 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	556	100.0	267	1	LPHUA1	apolipoprotein A-I
2	538	96.8	267	1	A26529	apolipoprotein A-I
3	538	96.8	267	2	JS0079	apolipoprotein A-I
4	449	80.8	266	1	LPDGA1	apolipoprotein A-I
5	430	77.3	266	1	LPRB12	apolipoprotein A-I
6	425	76.4	241	2	A24998	apolipoprotein A-I
7	418.5	75.3	265	1	LPRB1B	apolipoprotein A-I
8	414.5	74.6	265	2	A56858	apolipoprotein A-I
9	408.5	73.5	265	2	JT0672	apolipoprotein A-I
10	403.5	72.6	265	2	A46018	apolipoprotein A-I
11	400.5	72.0	164	2	S21830	apolipoprotein A-I
12	400.5	72.0	231	2	JQ0704	apolipoprotein A-I
13	400.5	72.0	264	2	S31394	apolipoprotein A-I
14	296	53.2	264	2	S22420	apolipoprotein A-I
15	290	52.2	262	2	JC1237	apolipoprotein A-I
16	273	49.1	259	2	A24700	apolipoprotein A-I
17	272	48.9	246	2	A61448	apolipoprotein A-I
18	255	45.9	264	1	LPCHA1	apolipoprotein A-I
19	237	42.6	264	2	JC5456	apolipoprotein A-I
20	125	22.5	396	1	LPHUA4	apolipoprotein A-I
21	119.5	21.5	429	2	S29565	apolipoprotein A-I
22	113.5	20.4	401	2	A47141	apolipoprotein A-I
23	111	20.0	391	1	LPRTA4	apolipoprotein A-I
24	111	20.0	1547	2	T28657	apolipoprotein A-I
25	109.5	19.7	258	2	JH0472	apolipoprotein A-I
26	102.5	18.4	391	2	B40892	apolipoprotein A-I
27	102.5	18.4	395	2	A40892	apolipoprotein A-I
28	102.5	18.4	399	2	C40892	apolipoprotein A-I
29	100.5	18.1	394	2	A25281	apolipoprotein A-I

30	99	17.8	317	2	A28792	apolipoprotein E p
31	99	17.8	317	2	S03185	apolipoprotein E p
32	98	17.6	316	2	JC6549	apolipoprotein E p
33	97.5	17.5	312	1	LPRTE	apolipoprotein E p
34	97.5	17.5	329	2	JC5566	apolipoprotein E p
35	95.5	17.2	311	2	JU0036	apolipoprotein E p
36	93.5	16.8	392	2	T38881	apolipoprotein E p
37	92.5	16.6	291	2	C60940	apolipoprotein E p
38	92	16.5	317	1	LPHUE	apolipoprotein E p
39	91.5	16.5	113	2	S66805	apolipoprotein E p
40	91	16.4	2155	2	C97523	hypothetical prote
41	91	16.4	2155	2	AD2742	hypothetical prote
42	90.5	16.3	3672	2	T23433	conserved hypothet
43	90.5	16.3	3704	2	T37316	hypothetical prote
44	90	16.2	316	2	S26478	probable laminin a
45	89.5	16.1	2241	2	S09811	apolipoprotein E - hypothetical prote

ALIGNMENTS

RESULT 1  
LPHUA1  
apolipoprotein A-I precursor [validated] - human  
N:Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C:Species: Homo sapiens (man)  
C:Date: 14-Nov-1983 #sequence, revision 14-Nov-1983 #text, change 08-Dec-2000  
C:Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010; 6197  
R:Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A:Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A:Reference number: A90947; MUID:85026665  
A:Accession: A90947  
A:Molecule type: DNA  
A:Residues: 1-267 <SEI>  
A:Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R:Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988  
A:Title: Sequence and expression of tangler apoA-I gene.  
A:Reference number: S02373; MUID:88196137  
A:Accession: S02373  
A:Molecule type: DNA  
A:Residues: 1-267 <MAK>  
A:Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
R:Shoulders, C.C.; Kornblith, A.R.; Munro, B.S.; Baralle, F.E.  
Nucleic Acids Res. 11, 2827-2837, 1983  
A:Title: Gene structure of human apolipoprotein A1.  
A:Reference number: A93465; MUID:83220822  
A:Accession: A93465  
A:Molecule type: DNA  
A:Residues: 1-267 <SHO>  
A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518;  
R:Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
A:Title: Isolation and characterization of the human apolipoprotein A-I gene.  
A:Reference number: A21147; MUID:84016011  
A:Accession: A21147  
A:Molecule type: DNA  
A:Residues: 1-267 <KAR>  
A:Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
R:Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.  
Nucleic Acids Res. 12, 3917-3932, 1984  
A:Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundan  
A:Reference number: A93519; MUID:84221405  
A:Accession: A93519  
A:Molecule type: mRNA  
A:Residues: 1-267 <SHA>  
A:Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753

A:Accession: B93519  
A:Molecule type: DNA  
A:Residues: 1-24 <SH2>  
R:Cheung, P.; Chan, L.  
Nucleic Acids Res. 11, 3703-3715, 1983  
A:Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
A:Reference number: A93472; MUID:83220772  
A:Accession: A93472  
A:Molecule type: mRNA  
A:Residues: 1-267 <CHE>  
A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R:Law, S.W.; Brewer Jr., H.B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
A:Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
A:Reference number: A94010; MUID:84119464  
A:Accession: A94010  
A:Molecule type: mRNA  
A:Residues: 1-267 <LAW>  
A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R:Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
A:Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
A:Reference number: A21118; MUID:83195100  
A:Accession: A21118  
A:Molecule type: protein  
A:Residues: 1-24 <ZAN>  
R:Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
A:Title: Human plasma proapoA-I: isolation and amino-terminal sequence.  
A:Reference number: A90112; MUID:83256553  
A:Accession: A90112  
A:Molecule type: protein  
A:Residues: 19-27 <BRE>  
R:Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
A:Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high density lipoproteins  
A:Reference number: A90209; MUID:78123731  
A:Accession: A90209  
A:Molecule type: protein  
A:Residues: 25-57, 'Q', '59-169, 'QQ', '172-267 <BR2>  
R:Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
J. Clin. Invest. 82, 803-807, 1988  
A:Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A I)  
A:Reference number: A30516; MUID:88331387  
A:Accession: A30516  
A:Molecule type: protein  
A:Residues: 25-56 <YUI>  
R:Nichols, W.C.; Dwulet, F.E.; Liepnies, J.; Benson, M.D.  
Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
A:Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid.  
A:Reference number: A31582; MUID:89050104  
A:Accession: A31582  
A:Molecule type: protein  
A:Residues: 25-49, 'R', '51-85, 'D', '87-107 <NIC>  
A:Note: variant sequence from patient with familial amyloidotic polyneuropathy type III  
R:Manjunath, P.; Marcel, Y.L.; Uma, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A.  
J. Biol. Chem. 264, 16853-16857, 1989  
A:Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
A:Reference number: A34409; MUID:89380318  
A:Accession: A34409  
A:Molecule type: protein  
A:Residues: 25-48 <MAN>  
R:Stoffel, W.; Binczek, E.  
Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
A:Title: Structural requirements of human preproapolipoprotein AI for translocation and secretion  
A:Reference number: S02737; MUID:89149957  
A:Accession: S02737  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-6, 'AV', '9, 'LV', '12-29 <STO>  
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by R:Stoffel, W.; Binczek, E.  
Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A:Title: Transient expression of wild type and mutant human apolipoprotein AI in COS cells  
A:Reference number: S16197; MUID:92029676  
A:Contents: annotation; extension of studies in reference S02737  
R:Stoffel, W.; Kruger, E.; Deutzmann, R.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
A:Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processed in vitro  
A:Reference number: A19913; MUID:83236195  
A:Accession: B19913  
A:Molecule type: protein  
A:Residues: 1-6, 'X', '8-13, 'XXX', '17-18, 'XX', '21, 'X', '23-25, 'X', '27-29 <ST2>  
R:Elmholm, C.; Bozas, S.E.; Tenkanen, H.; Kirsbaum, L.; Metso, J.; Murphy, B.; Walke  
Biochim. Biophys. Acta 1086, 255-260, 1991  
A:Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
A:Reference number: A56815; MUID:92075698  
A:Accession: A56815  
A:Molecule type: protein  
A:Residues: 25-31, 'P', '33 <EHN>  
A:Experimental source: serum  
A:Note: sequence extracted from NCBI backbone (NCBIP:69759)  
A:Note: 32-Tyr was also found  
R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
Biochemistry 33, 1988-1993, 1994  
A:Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins  
A:Reference number: A54223; MUID:94162201  
A:Accession: A54223  
A:Molecule type: protein  
A:Residues: 25-39 <KUN>  
R:Mogulovsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.; DNA 8, 429-436, 1989  
A:Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: purification and characterization  
A:Reference number: I39476; MUID:89377481  
A:Accession: I39476  
A:Molecule type: mRNA  
A:Residues: 19-267 <RES>  
A:Cross-references: GB:M29068; NID:q178774; PIDN:AAA51747.1; PID:q178775  
R:Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B.  
J. Biol. Chem. 263, 18530-18536, 1988  
A:Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by the APOA1 gene  
A:Reference number: I39475; MUID:89054040  
A:Accession: I39475  
A:Molecule type: DNA  
A:Residues: 1-14 <RE2>  
A:Cross-references: GB:J04066; NID:q178763; PIDN:AAA51746.1; PID:q553183  
R:Breslow, J.L.  
Ann. Rev. Biochem. 54, 699-727, 1985  
A:Title: Human apolipoprotein molecular biology and genetic variation.  
A:Reference number: A90042; MUID:85278004  
A:Contents: annotation; review of sequences, variants and gene location  
R:Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B.  
J. Biol. Chem. 261, 3911-3914, 1986  
A:Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation  
A:Reference number: A92577; MUID:86140194  
A:Contents: annotation; acylation with palmitate  
A:Note: an undetermined serine or threonine is acylated by fatty acid; the acylating agent is not determined  
R:Law, S.W.; Brewer, H.B.  
J. Biol. Chem. 260, 12810-12814, 1985  
A:Title: Tangier disease: The complete mRNA sequence encoding for preproapoA-I.  
A:Reference number: I55236; MUID:86008382  
A:Accession: I55236  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-143, 'D', '145-267 <RE3>  
A:Cross-references: GB:M11791; NID:q178776; PIDN:AAA35545.1; PID:q178777  
C:Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestine (HDL) in plasma.  
C:Genetics:  
A:Gene: GDB:AP0A1  
A:Cross-references: GDB:119684; OMIM:107680  
A:Map position: 11q23.3-11q23.3  
A:Introns: 15/1; 67/2  
C:Function:  
A:Description: participates in the reverse transport of cholesterol from tissues to the liver and small intestine; noncovalently binds and stabilizes prostacyclin (PGI-2) and acyltransferase (LCAT);





A:Molecule type: protein  
A:Residues: 25-56,'Z';261-262,'A' <NAK>  
C:Superfamily: apolipoprotein A-I  
C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid transport; 14  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 80.8%; Score 449; DB 1; Length 266;  
Best Local Similarity 80.4%; Pred. No. 1.9e-29;  
Matches 90; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 QKLHLEQKLSPLGEEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGARLAE 60  
Db 155 QKLQLEQKLSPLAEELRDSARTHVDTLRKLAAPYSDDLRERLAARLEALKENGASLAE 214

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112  
Db 215 YHARASEQLSALGEKARPALEDLRQGLLPVLESFKVSLAIDEATKKLNAQ 266

RESULT 5  
LPRB12  
apolipoprotein A-I precursor (clone 22ap AI) - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S06064

R;Paraskevopoulou, T.B.; Kritis, A.; Zannis, V.  
submitted to the EMBL Data Library, July 1989

A:Reference number: S06064

A:Accession: S06064  
A:Molecule type: mRNA

A:Residues: 1-266 <PAR>

A:Cross-references: EMBL:X15908; NID:g1457; PIDN:CAA34024.1; PID:g1458

C:Comment: This protein is synthesized in the small intestine.

C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe

F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 77.3%; Score 430; DB 1; Length 266;  
Best Local Similarity 78.6%; Pred. No. 6.8e-28;  
Matches 88; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 1 QKLHLEQKLSPLGEEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGARLAE 60  
Db 155 QKLTELQEKLSPLAEELRDSARTHVDTLRKLAAPYSNELQRLAARLESIKEGGASLAE 214

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112  
Db 215 YQAKAREHLSVLSSEKARPALEDLRQGLLPVLESFKASVQNVNDEATKKLTQ 266

RESULT 6  
A24998

apolipoprotein A-I - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 31-Dec-1993

C:Accession: A24998

R;Yang, C.; Yang, T.; Pownall, H.J.; Gotto Jr., A.M.

Eur. J. Biochem. 160, 427-431, 1986

A:Title: The primary structure of apolipoprotein A-I from rabbit high-density lipoprotei  
A:Reference number: A24998; MUID:87030294

A:Accession: A24998

A:Molecule type: protein

A:Residues: 1-241 <YAN>

C:Superfamily: apolipoprotein A-I

C:Keywords: HDL; lipid binding; lipoprotein

Query Match 76.4%; Score 425; DB 2; Length 241;  
Best Local Similarity 77.7%; Pred. No. 1.5e-27;  
Matches 87; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

QY 1 QKLHLEQKLSPLGEEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGARLAE 60  
Db 130 QKLTELQEKLSPLAEELRDSARTHVDTLRKLAAPYSQELQRLAARLESIKEGGASLAE 189

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112  
Db 190 YQAKAREHLSVLSSEKARPALEDLRQGLLPVLESFKASVQNVNDEATKKLTQ 241

RESULT 7  
LPRB1B  
apolipoprotein A-I precursor (clone PRBA-502) - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999

C:Accession: S00230; S20557

R;Pan, T.C.; Hao, Q.L.; Yamin, T.T.; Dai, P.H.; Chen, B.S.; Chen, S.L.; Kroon, P.A.;  
Eur. J. Biochem. 170, 99-104, 1987

A:Title: Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit apolipoprotein

A:Reference number: S00230; MUID:88082865

A:Accession: S00230

A:Molecule type: mRNA

A:Residues: 1-265 <PAN>

A:Cross-references: EMBL:X06658; NID:g1461; PIDN:CAA29857.1; PID:g1462

A:Note: the authors translated the codon AGC for residue 174 as Arg

A:Accession: S20557

A:Molecule type: DNA

A:Residues: 1-17,'R',19-44,'T',46-122,'V',124-146,'V',148-265 <PAN2>

A:Cross-references: EMBL:X06659; NID:g1459; PIDN:CAA29858.1; PID:g1460

C:Comment: This protein is synthesized in the small intestine.

C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r

F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 75.3%; Score 418.5; DB 1; Length 265;  
Best Local Similarity 77.7%; Pred. No. 5.8e-27;  
Matches 87; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 QKLHLEQKLSPLGEEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGARLAE 60  
Db 155 QKLTELQEKLSPLAEELRDSARTHVDTLRKLAAPYSNEL-QRLAARLESIKEGGAKLAE 213

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112  
Db 214 YQAKAREHLSVLSSEKARPALEDLRQGLLPVLESFKASVQNVNDEATKKLTQ 265

RESULT 8  
A56858

apolipoprotein A-I precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 05-Jan-1996 #sequence\_revision 23-Aug-1997 #text\_change 13-Aug-1999

C:Accession: I45853; A56858; A34649

R;O'Huigin, C.; Chan, L.; Li, W.

Mol. Biol. Evol. 7, 327-339, 1990

A:Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolu

A:Reference number: I45853; MUID:90348478

A:Accession: I45853

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-265 <OXH>

A:Cross-references: GB:M35870; NID:g162677; PIDN:AAA30381.1; PID:g162678

R;Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Auboiron, S.; Bauchart, D.; Chapman, M.J.;

Biochim. Biophys. Acta 1123, 145-150, 1992  
A;Title: Plasma lipid transport in the prerinant calf, Bos spp: primary structure of h  
A;Reference number: A56858; MUID:92153895  
A;Accession: A56858  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 19-184,'QL','187-265 <SPA>  
A;Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
A;Note: sequence extracted from NCBI backbone (NCBIP:83520)  
R;Aubolton, S.; Sparrow, D.A.; Beaubatie, L.; Bauchart, D.; Sparrow, J.T.; Laplaud, P.M.  
Biochem. Biophys. Res. Commun. 166, 833-839, 1990  
A;Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasma h  
A;Reference number: A34649; MUID:90147795  
A;Accession: A34649  
A;Molecule type: protein  
A;Residues: 25-70 <AUB>  
A;Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
C;Superfamily: apolipoprotein A-I  
C;Keywords: lipid binding, lipoprotein  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match	74.6%;	Score 414.5;	DB 2;	Length 265;
Best Local Similarity	74.1%;	Pred. No. 1.2e-26;		
Matches	83;	Conservative	13;	Mismatches 15;
			Indels	1;
			Gaps	1;

[illegible]

RESULT 9  
JT0672

apolipoprotein A-I - pig  
N:Alternate names: apo-A-I  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 11-May-2000  
C:Accession: J0672; P00471; A05311  
R:Trieu, V.N.; Patel, B.; Zhan, R.; Black, D.D.  
Gene 134, 267-270, 1993  
A:Title: Sequence of the porcine apoA-I gene.  
A:Reference number: J0672; MUID:94085789  
A:Accession: J0672  
A:Molecule type: DNA  
A:Residues: 1-265 <TRI>  
A:Cross-references: EMBL:Z14124; NID:g1893  
A:Note: this translation is not annotated in GenBank entry SSAPOAIG, release 111.0; the  
R:Trieu, V.N.; Hasler-Rapacz, J.; Rapacz, J.; Black, D.D.  
Gene 123, 173-179, 1993  
A:Title: Sequences and expression of the porcine apolipoprotein A-I and C-III mRNAs.  
A:Reference number: P00471; MUID:93154581  
A:Accession: P00471  
A:Molecule type: mRNA  
A:Residues: 105-265 <TR2>  
A:Experimental source: liver  
R:Mahley, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.  
Biochemistry 15, 1928-1933, 1976  
A:Title: Characterization of the plasma lipoproteins and apoproteins of the Erythrocebus  
A:Reference number: A90395; MUID:76184721  
A:Accession: A05311  
A:Molecule type: protein  
A:Residues: 25-34 <MAH>  
C:Comment: This protein is the major apolipoprotein of high-density lipoprotein and serves  
C:Genetics:  
A:Gene: apoA-I  
A:Introns: 15/1, 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid metabolism; lip

F; 99/Region: ochre stop codon

Query Match	73.5%;	Score 408.5;	DB 2;	Length 265;
Best Local Similarity	74.1%;	Pred. No. 3.7e-26;		
Matches 83; Conservative	12;	Mismatches 16;	Indels 1;	Gaps 1;

[illegible]

RESULT 10  
A46018

apolipoprotein AI - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
 C:Accession: A46018  
 R:Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.  
 Genomics 15, 643-652, 1993  
 A:Title: Characterization of the apolipoprotein AI and CIII genes in the domestic pig  
 A:Reference number: A46018; MUID:93224154  
 A:Accession: A46018  
 A:Status: preliminary  
 A:Molecule type: DNA; protein  
 A:Residues: 1-265 <BIR>  
 A:Cross-references: GB:I00626; NID:g164358; PIDN:AAA30992.1; PID:g164359  
 A:Note: sequence extracted from NCBI backbone (NCBIN:129509, NCBI:P:129511)  
 C:Superfamily: apolipoprotein A-I

Query Match	72.6%;	Score 403.5;	DB 2;	Length 265;
Best Local Similarity	74.1%;	Pred. No. 9.5e-26;		
Matches 83;	Conservative 11;	Mismatches 17;	Indels 1;	Gaps 1;

[illegible]

RESULT 11  
S21830

apolipoprotein A-I - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 20-Feb-1995 #sequence\_revision 22-May-1998 #text\_change 13-Aug-1999  
 C:Accession: S21830  
 R:Trieu, V.N.; Rohwer-Nuttler, P.; Black, D.D.  
 submitted to the EMBL Data Library, May 1991  
 A:Description: Sequence and developmental expression of porcine ApoA-I mRNA.  
 A:Reference number: S21830  
 A:Accession: S21830  
 A:Molecule type: mRNA  
 A:Residues: 1-164 <TRI>  
 A:Cross-references: EMBL:X59414, NID:g1891, PIDN:CAA42050.1, PID:g1892  
 A:Experimental source: liver  
 C:Superfamily: apolipoprotein A-I  
 C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid metabolism;

Query Match	72.0%;	Score 400.5;	DB 2;	Length 164;
Best Local Similarity	73.2%;	Pred. No. 1e-25;		
Matches 82; Conservative	12;	Mismatches 17;	Indels 1;	Gaps 1;

01 OKLHETOELISPLGEEMDRARAHVDALRTHLAPYSDELKORLARLEALKENGARLAE 60  
02  
03  
04  
05  
06  
07  
08  
09  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
00

Db 54 QKVQLQEKLSPLAEELRSRLRAHVEALRQHVAPYSDDLQRMARFEALKEGGSLAE 112  
 QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLNQ 112  
 Db 113 YQAKAQEQLKGLGERAKPALEDLRQGLLPVLENLKVSTILAIDEASKKLNAQ 164

# RESULT 12

JQ0704  
 apolipoprotein A-I - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 28-Oct-1994  
 C:Accession: JQ0704  
 R;Weiler-Guettler, H.; Sommerfeldt, M.; Papandrikopoulou, A.; Mischek, U.; Bonitz, D.; F. J. Neurochem. 54, 444-450, 1990  
 A:Title: Synthesis of apolipoprotein A-I in pig brain microvascular endothelial cells.  
 A:Reference number: JQ0704; MUID:90132667  
 A:Accession: JQ0704  
 A:Molecule type: mRNA  
 A:Residues: 1-231 <WEI>  
 A>Note: the authors translated the codon CAG for residue 124 as His and GAC for residue C;Superfamily: apolipoprotein A-I  
 C;Keywords: cholesterol metabolism; HDL; lipid binding; lipid transport; lipoprotein; pl

Query Match 72.0%; Score 400.5; DB 2; Length 231;  
 Best Local Similarity 74.1%; Pred. No. 1.4e-25;  
 Matches 83; Conservative 10; Mismatches 18; Indels 1; Gaps 1;

QY 1 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKEGARGLA 60  
 Db 121 QKVQLQEKLSPLAEELRDLRAHVAALRQHVAPYSDDLQRMARFEALKE-GGDSLAE 179  
 QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLNQ 112  
 Db 180 YQAKAQEQLKGLGERAKPALEDLRQGLLPVLENLKVSTILAIDEASKKLNAQ 231

# RESULT 13

S31394  
 apolipoprotein A-I - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
 C:Accession: S31394  
 R;Moechel, B.; Flach, R.; Weiss, B.; Weiler-Guettler, H.; Frey, A.; Zinke, H.; Gassen, H. submitted to the EMBL Data Library, November 1992  
 A:Description: Genomic organization of the porcine apolipoprotein AI gene and study of g  
 A:Reference number: S31394  
 A:Accession: S31394  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-264 <MOE>  
 A:Cross-references: EMBL:X69477; NID:g1889; PIDN:CAA49234.1; PID:g1390  
 C;Superfamily: apolipoprotein A-I

Query Match 72.0%; Score 400.5; DB 2; Length 264;  
 Best Local Similarity 74.1%; Pred. No. 1.7e-25;  
 Matches 83; Conservative 10; Mismatches 18; Indels 1; Gaps 1;

QY 1 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKEGARGLA 60  
 Db 154 QKVQLQEKLSPLAEELRDLRAHVAALRQHVAPYSDDLQRMARFEALKE-GGDSLAE 212  
 QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLNQ 112  
 Db 213 YQAKAQEQLKGLGERAKPALEDLRQGLLPVLENLKVSTILAIDEASKKLNAQ 264

RESULT 14  
 S22420  
 apolipoprotein A-I precursor - mouse  
 C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
 C:Accession: S22420; S22421; A44364  
 R;Stoffel, W.; Mueller, R.; Binczek, E.; Hofmann, K.  
 Biol. Chem. Hoppe-Seyler 373, 187-193, 1992  
 A:Title: Mouse apolipoprotein AI. cDNA-derived primary structure, gene organisation a  
 A:Reference number: S22420; MUID:92281682  
 A:Accession: S22420

A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-264 <STO>  
 A:Cross-references: EMBL:X64262; NID:g50014; PIDN:CAA45560.1; PID:g50015  
 A:Accession: S22421  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-264 <ST2>  
 A:Cross-references: EMBL:X64263; NID:g50020; PIDN:CAA45561.1; PID:g50021  
 R;Januzzi, J.L.; Azrolan, N.; O'Connell, A.; Aalto-Setälä, K.; Breslow, J.L.  
 Genomics 14, 1081-1088, 1992  
 A:Title: Characterization of the mouse apolipoprotein Apoa-1/Apoc-3 gene locus: genom  
 A:Reference number: A44364; MUID:93122774  
 A:Accession: A44364  
 A>Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-264 <JAN>  
 A>Note: sequence extracted from NCBI backbone (NCBIN:122400, NCBI:P:122407)  
 C:Genetics:  
 A:Introns: 15/1; 66/2  
 C;Superfamily: apolipoprotein A-I

Query Match 53.2%; Score 296; DB 2; Length 264;  
 Best Local Similarity 53.6%; Pred. No. 5e-17;  
 Matches 60; Conservative 18; Mismatches 32; Indels 2; Gaps 1;

QY 1 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKEGARGLA 60  
 Db 155 QKLQELQGRLSPVAEFRDRMRTHVDSLRTQLAPHSEQMRRESLAQRRLAELKSN--PTLINE 212  
 QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLNQ 112  
 Db 213 YHTRAKTHLKTGLGERAKPALEDLRHSLMPMLFTLTKAOSVIDKASETLTAO 264

# RESULT 15

JC1237  
 apolipoprotein A-I precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-1993  
 C:Accession: JC1237  
 R;Boyle, T.P.; Marotti, K.R.  
 Gene 117, 243-247, 1992  
 A:Title: Structure of the murine gene encoding apolipoprotein A-I.  
 A:Reference number: JC1237; MUID:92347700  
 A:Accession: JC1237  
 A:Molecule type: DNA  
 A:Residues: 1-262 <BOY>  
 A:Cross-references: GB:M77801  
 C:Genetics:  
 A:Gene: Apoa-I  
 A:Introns: 15/1; 66/2  
 C;Superfamily: apolipoprotein A-I  
 C;Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; l  
 F;1-24/Domain: signal sequence #status predicted <SIG>  
 F;25-262/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 52.2%; Score 290; DB 2; Length 262;  
 Best Local Similarity 57.3%; Pred. No. 1.5e-16;  
 Matches 59; Conservative 14; Mismatches 28; Indels 2; Gaps 1;

QY 1 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKEGARGLA 60  
 Db 155 QKLQELQGRLSPVAEFRDRMRTHVDSLRTQLAPHSEQMRRESLAQRRLAELKSN--PTLINE 212



Sun Sep 22 12:10:12 2002

us-09-803-918a-2\_copy\_156\_267.rpr

Page 7

```
QY      61 YHAKATEHISTLSEKAKPALEDIRQGLLPVLESEKVSFLSALE 103
      || : | || | | | | | | | | | | | | | | | | | | | | | | | |
Db      213 YHTRAKTHLTKLGEKARPALEDLRHSLMPMLETLTKKAQSVIE 255
```

Search completed: September 22, 2002, 12:06:29  
Job time: 324 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:22:18 ; Search time 35.02 Seconds  
(without alignments)  
123.832 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_156\_267

Perfect score: 556  
Sequence: 1 QKLHELOEKLSPGEMRDR.....SFKVSFLSALLEEYTKKLNTQ 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	556	100.0	267	1 APAL_HUMAN	P02647 homo sapien
2	538	96.8	267	1 APAL_MACFA	P15568 macaca fasc
3	449	80.8	266	1 APAL_CANFA	P02648 canis fami1
4	430	77.3	266	1 APAL_RABIT	P09809 oryctolagus
5	409.5	73.7	265	1 APAL_BOVIN	P15497 bos taurus
6	408.5	73.5	265	1 APAL_PIG	P18648 sus scrofa
7	355	63.8	265	1 APAL_TUPGB	O18759 tupai glis
8	296	53.2	264	1 APAL_MOUSE	Q00623 mus musculu
9	272	48.9	264	1 APAL_ANAPL	O42296 anas platyr
10	266	47.8	259	1 APAL_RAT	P04639 rattus norv
11	255	45.9	264	1 APAL_CHICK	P08250 gallus gall
12	237	42.6	264	1 APAL_COTJA	P32918 coturnix co
13	124	22.3	396	1 APAL_HUMAN	P06727 homo sapien
14	122	21.9	262	1 APAL_BRARE	O42363 brachydanio
15	119.5	21.5	382	1 APAL_PIG	O46409 sus scrofa
16	119.5	21.5	429	1 APAL_MACFA	P33621 macaca fasc
17	116	20.9	262	1 AP12_ONCMY	O57524 oncorhynchu
18	115	20.7	262	1 AP11_ONCMY	O57523 oncorhynchu
19	113.5	20.4	401	1 AP44_PAPAN	Q28758 papio anubi
20	111.5	20.1	262	1 APAL_SALTR	Q91488 salmo trutt
21	107.5	19.3	260	1 APAL_SPAUV	O42175 sparus aura
22	105.5	19.0	258	1 APAL_SALSA	P27007 salmo salar
23	103	18.5	391	1 AP44_MOUSE	P02651 rattus norv
24	102.5	18.4	395	1 AP44_MOUSE	P06728 mus musculu
25	99	17.8	317	1 APAL_MACFA	P10517 macaca fasc
26	99	17.8	317	1 APAL_PAPAN	P05770 papio anubi
27	97.5	17.5	312	1 APAL_PAPAN	P02650 rattus norv
28	97	17.4	107	1 APAL_SAISC	Q28995 salmimiri sci
29	95.5	17.2	311	1 APAL_MOUSE	P08226 mus musculu
30	93.5	16.8	392	1 YAY1_SCHPO	O10209 schizosacch
31	93	16.7	107	1 APAL_MACMU	Q28502 macaca mula
32	92.5	16.6	305	1 APAL_CANFA	P18649 canis fami1
33	92	16.5	281	1 APAL_BRARE	O42364 brachydanio

34	92	16.5	317	1	APAL_HUMAN	P02649 homo sapien
35	90.5	16.3	3672	1	LML2_CAEL	Q21313 caenorhabdi
36	90	16.2	316	1	APAL_BOVIN	Q03247 bos taurus
37	89.5	16.1	2241	1	TEGU_HCMVA	P16785 human cytom
38	87	15.6	317	1	APAL_PIG	P18650 sus scrofa
39	85	15.3	311	1	APAL_RABIT	P18287 oryctolagus
40	84.5	15.2	1935	1	MYH7_HUMAN	P12883 homo sapien
41	84	15.1	857	1	MEDE_GIALA	Q08014 giardia lam
42	84	15.1	2411	1	MYSA_DROME	P05661 drosophila
43	83	14.9	298	1	APAL_CAVPO	P23529 cavia porce
44	83	14.9	512	1	K2C5_XENLA	P16878 xenopus lae
45	83	14.9	1184	1	BIMC_EMENTI	P17120 emerice11a

## ALIGNMENTS

RESULT	1	STANDARD;	PRT;	267 AA.
APAL_HUMAN				
ID	APAL_HUMAN			
AC	P02647;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).			
GN	APOL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84221405; Pubmed=6328445;			
RA	Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,			
RA	Baralle F.E.;			
RT	"Human apolipoproteins AI, AII, CII and CIIT. cDNA sequences and mRNA			
RT	abundance."			
RL	Nucleic Acids Res. 12:3917-3932(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8502665; Pubmed=6207999;			
RA	Selhamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;			
RT	"Isolation and DNA sequence of full-length cDNA and of the entire			
RT	gene for human apolipoprotein AI -- discovery of a new genetic			
RT	polymorphism in the apo AI gene."			
RL	DNA 3:309-317(1984).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83220822; Pubmed=6406984;			
RA	Shoulders C.C., Kornblitt A.R., Munro B.S., Baralle F.E.;			
RT	"Gene structure of human apolipoprotein AI."			
RL	Nucleic Acids Res. 11:2827-2837(1983).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83220772; Pubmed=6304641;			
RA	Cheung P., Chan L.;			
RT	"Nucleotide sequence of cloned cDNA of human apolipoprotein A-I."			
RL	Nucleic Acids Res. 11:3703-3715(1983).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84119464; Pubmed=6198645;			
RA	Law S.W., Brewer H.B. Jr.;			
RT	"Nucleotide sequence and the encoded amino acids of human			
RT	apolipoprotein A-I mRNA."			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86008382; Pubmed=2995392;			
RA	Law S.W., Brewer H.B. Jr.;			
RT	"Tangier disease. The complete mRNA sequence encoding for			
RT	preproapo-A-I."			
RL	J. Biol. Chem. 260:12810-12814(1985).			
RN	[7]			

RP SEQUENCE FROM N.A.  
RX MEDLINE=84016011; PubMed=6413973;  
RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
RT "Isolation and characterization of the human apolipoprotein A-I  
gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
RN [18]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89377481; PubMed=2673706;  
RA Mogullevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,  
RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
RA Holmquist L., Carlson L.A., Bollen A.;  
RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
coli: purification and biochemical characterization.";  
RL DNA 8:429-436(1989).  
RN [19]  
RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
RX MEDLINE=8196137; PubMed=3129297;  
RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
RA Zannis V.I.;  
RT "Sequence and expression of Tangier apoA-I gene.";  
RL Eur. J. Biochem. 173:465-471(1988).  
RN [10]  
RP SEQUENCE OF 118-267 FROM N.A.  
RX MEDLINE=83091059; PubMed=6294659;  
RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
RT "Isolation and characterization of cDNA clones for human  
apolipoprotein A-I.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
RN [11]  
RP SEQUENCE OF 19-27.  
RX MEDLINE=83256553; PubMed=6409108;  
RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
RA Light J.A.;  
RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
RN [12]  
RP SEQUENCE OF 25-267.  
RX MEDLINE=78123731; PubMed=204308;  
RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
RA Bronzert T.J.;  
RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated  
from high density lipoproteins.";  
RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
RN [13]  
RP SEQUENCE OF 25-267.  
RX MEDLINE=75133493; PubMed=164450;  
RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;  
RT "The primary structure of human plasma high density apolipoprotein  
glutamine I (ApoA-I). II. The amino acid sequence and alignment of  
cyanogen bromide fragments IV, III, and I.";  
RL J. Biol. Chem. 250:2725-2738(1975).  
RN [14]  
RP SEQUENCE OF 25-56.  
RX MEDLINE=88331387; PubMed=3047170;  
RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
A-I (Apo A-I). A novel function of Apo A-I.";  
RL J. Clin. Invest. 82:803-807(1988).  
RN [15]  
RP SEQUENCE OF 25-48.  
RX MEDLINE=89380318; PubMed=2506184;  
RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
RA Chappelaine A.;  
RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
proteins.";  
RL J. Biol. Chem. 264:16853-16857(1989).  
RN [16]  
RP SEQUENCE OF 25-43.  
RX MEDLINE=88070603; PubMed=3120314;  
RA Prioli R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
RA Pereira M.E.A.;

RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
neuraminidase, to high-density lipoprotein.";  
RL Science 238:1417-1419(1987).  
RN [17]  
RP SEQUENCE OF 25-42.  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [18]  
RP PALMITOYLATION.  
RX MEDLINE=86140194; PubMed=3005308;  
RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
RT "Human apolipoprotein A-I. Post-translational modification by fatty  
acid acylation.";  
RL J. Biol. Chem. 261:3911-3914(1986).  
RN [19]  
RP PROCESSING.  
RX MEDLINE=83195100; PubMed=6405383;  
RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
RA Breslow J.L.;  
RT "Intracellular and extracellular processing of human apolipoprotein  
A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
RN [20]  
RP STRUCTURE BY NMR OF 190-209.  
RX MEDLINE=96270776; PubMed=8664326;  
RA Wang G., Treleaven W.D., Cushley R.J.;  
RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
and CD. Evidence for specific peptide-SDS interactions.";  
RL Biochim. Biophys. Acta 1301:174-184(1996).  
RN [21]  
RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
RX MEDLINE=98024124; PubMed=9356442;  
RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
lipid-bound conformation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
RN [22]  
RP VARIANT MILANO.  
RX MEDLINE=83109095; PubMed=6401735;  
RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,  
RA Franceschini G., Sirtori C.R.;  
RT "Apolipoprotein A-I-Milano. Detection of normal A-I in affected  
subjects and evidence for a cysteine for arginine substitution in the  
variant A-I.";  
RL J. Biol. Chem. 258:2508-2513(1983).  
RN [23]  
RP VARIANT TANGIER.  
RX MEDLINE=83300108; PubMed=6412234;  
RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;  
RT "Tangier disease: defective recombination of a specific Tangier  
apolipoprotein A-I isoform (pro-apo A-I) with high density  
lipoproteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).  
RN [24]  
RP VARIANT NORWAY.  
RX MEDLINE=84289383; PubMed=6432779;  
RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,  
RA Utermann G., Haas J., Steinmetz A., Menzel H.J., Assmann G.;  
RT "Abnormal lecithin:cholesterol acyltransferase activation by a human  
apolipoprotein A-I variant in which a single lysine residue is  
deleted.";  
RL J. Biol. Chem. 259:10063-10070(1984).  
RN [25]  
RP SEQUENCE OF 25-107 (VARIANT IOWA).  
RX MEDLINE=89050104; PubMed=3142462;  
RA Nichols W.C., Dwulet F.E., Diepnieks J., Benson M.D.;  
RT "Variant apolipoprotein AI as a major constituent of a human  
hereditary amyloid.";



```

Query Match          100.0%; Score 556; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-39;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OKLHELOEKLSPGEMDRARAHDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60
    |||||||
Db 156 OKLHELOEKLSPGEMDRARAHDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 215

OY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112
    |||||||
Db 216 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 267

RESULT 2
APAL_MACFA STANDARD; PRT; 267 AA.
AC P15568; P17929;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OC Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541, 9557;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.fascicularis;
RX MEDLINE=87191989; PubMed=3106152;
RA Polites H.G., Melchior G.W., Castle C.K., Marotti K.R.;
RT "The primary structure of cynomolgus monkey apolipoprotein A-I
    deduced from the cDNA sequence: comparison to the human sequence.";
RL Gene 49:103-110(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.fascicularis;
RX MEDLINE=92305062; PubMed=1610902;
RA Murray R.W., Marotti K.R.;
RT "Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene
    and corresponding flanking regions.";
RL Biochim. Biophys. Acta 1131:207-210(1992).
RN [3]
RP SEQUENCE OF 25-48.
RC SPECIES=M.fascicularis;
RX MEDLINE=87185451; PubMed=3105581;
RA Herbert P.N., Bausserman L.L., Lynch K.M., Saritelli A.L.,
    Kantor M.A., Nicolosi R.J., Shulman R.S.;
RT "Homologues of the human C and A apolipoproteins in the Macaca
    fascicularis (cynomolgus) monkey.";
RL Biochemistry 26:1457-1463(1987).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=P.hamadryas; TISSUE=Liver;
RX MEDLINE=89232739; PubMed=2907746;
RA Hixson J.E., Borenstein S., Cox L.A., Rainwater D.L., Vandenberg J.L.;
RT "The baboon gene for apolipoprotein A-I: characterization of a cDNA
    clone and identification of DNA polymorphisms for genetic studies of
    cholesterol metabolism.";
RL Gene 74:483-490(1988).
RN [5]
RP SEQUENCE OF 1-10 FROM N.A.
RC SPECIES=M.fascicularis;
RA Sorci-Thomas M.;
RL Submitted (OCT-1991) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.

```

```

CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15411; AAA36834.1; -
DR EMBL; M83242; AAA36832.1; -
DR EMBL; M35634; AAA35380.1; -
DR EMBL; M69223; AAA36831.1; -
DR PIR; A26529; A26529.
DR PIR; A26627; A26627.
DR PIR; JS0079; JS0079.
DR PIR; S23135; S23135.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 267
FT DOMAIN 68 267
FT REPEAT 68 89
FT REPEAT 90 111
FT REPEAT 112 122
FT REPEAT 123 144
FT REPEAT 145 166
FT REPEAT 167 188
FT REPEAT 189 210
FT REPEAT 211 232
FT REPEAT 233 243
FT REPEAT 244 267
FT REPEAT 267 267
FT CONFLICT 13 13
FT SEQUENCE 267 AA; 30735 MW; 869955C024088E21 CRC64;
L -> P (IN REF. 1).

Query Match          96.8%; Score 538; DB 1; Length 267;
Best Local Similarity 96.4%; Pred. No. 5e-38;
Matches 108; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 OKLHELOEKLSPGEMDRARAHDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60
    |||||||
Db 156 OKLHELOEKLSPGEMDRARAHDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 215

OY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112
    |||||||
Db 216 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 267

RESULT 3
APAL_CANFA STANDARD; PRT; 266 AA.
AC P02648;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90132271; PubMed=2515239;
RA Luo C.-C., Li W.-H., Chan L.;

```

RT "Structure and expression of dog apolipoprotein A-I, E, and C-I  
 RT mRNAs: implications for the evolution and functional constraints of  
 RT apolipoprotein structure.";  
 RL J. Lipid Res. 30:1735-1746(1989).  
 RN [2]  
 RP SEQUENCE OF 25-266.  
 RX MEDLINE=82142425; PubMed=6801039;  
 RA Chung H., Randolph A., Reardon I., Heinrichson R.L.;  
 RT "The covalent structure of apolipoprotein A-I from canine high  
 RT density lipoproteins.";  
 RL J. Biol. Chem. 257:2961-2967(1982).  
 RN [3]  
 RP SEQUENCE OF 25-57 AND 262-265.  
 RX MEDLINE=76210910; PubMed=179887;  
 RA Nakai T., Whayne T.F., Tang J.;  
 RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein  
 RT A-I.";  
 RL FEBS Lett. 64:409-411(1976).  
 RN [4]  
 RP SEQUENCE OF 25-37.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 RT dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL EFFLUX FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS. SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 DR PIR: A03092; LPPGAL.  
 DR HSSP: P02647; 1ODR.  
 DR HSC-2DPAGE; P02648; DOG.  
 DR InterPro: IPR000074; Apolipoprotein.  
 DR Pfam: PF01442; Apolipoprotein; 1.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT PROPEP 19 24  
 FT CHAIN 25 266  
 FT DOMAIN 67 266  
 FT REPEAT 67 88  
 FT REPEAT 89 110  
 FT REPEAT 111 121  
 FT REPEAT 122 143  
 FT REPEAT 144 165  
 FT REPEAT 166 187  
 FT REPEAT 188 209  
 FT REPEAT 210 231  
 FT REPEAT 232 242  
 FT REPEAT 243 266  
 FT REPEAT 266 266  
 FT CONFLICT 168 168 A -> G (IN REF. 2).  
 FT CONFLICT 202 202 E -> Q (IN REF. 2).  
 FT CONFLICT 235 235 E -> Q (IN REF. 2).  
 FT CONFLICT 264 266 NAQ -> A (IN REF. 3).  
 SQ SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;

Query Match 80.8%; Score 449; DB 1; Length 266;  
 Best Local Similarity 80.4%; Pred. No. 1.2e-30;  
 Matches 90; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 OKLHELOEKLSPLGEMRDRAHVDALRTHLAPYSDELQRLAARLEALKENGARLAE 60  
 DB 155 OKLOEIOEKLSPLAEELRDARHVDALRAQLAPYSDLRERLAARLEALKENGASIAE 214  
 QY 61 YHAKATEHLSTLSEKAPPALEDLROGLLPVLESFKVSLAEYEYTKKLNTQ 112  
 DB 215 YHARASEQLSALGEKARPALEDLROGLLPVLESFKVSLAIDEATKKLNAQ 266

RESULT 4  
 ID APAL\_RABIT STANDARD; PRT; 266 AA.  
 AC P09809;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=22AP AI; TISSUE=Small intestine;  
 RA Paraskevopoulou T.B., Kritlis A., Zannis V.I.;  
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Intestine;  
 RX MEDLINE=88082866; PubMed=3121329;  
 RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,  
 RA Kroon P.A., Chao Y.S.;  
 RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit  
 RT apolipoprotein A-I is synthesized in the intestine but not in the  
 RT liver.";  
 RL Eur. J. Biochem. 170:99-104(1987).  
 RN [3]  
 RP SEQUENCE OF 25-266.  
 RX MEDLINE=87030294; PubMed=3095115;  
 RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;  
 RT "The primary structure of apolipoprotein A-I from rabbit high-density  
 RT lipoprotein.";  
 RL Eur. J. Biochem. 160:427-431(1986).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL EFFLUX FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X15908; CAA34024.1; -;  
 DR EMBL: X06658; CAA29857.1; -;  
 DR EMBL: X06659; CAA29858.1; -;  
 DR PIR: S06064; LPRB1Z.  
 DR PIR: S00230; LPRB1B.  
 DR PIR: A24998; A24998.  
 DR HSSP: P02647; 1AV1.  
 DR InterPro: IPR000074; Apolipoprotein.  
 DR Pfam: PF01442; Apolipoprotein; 1.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 266  
 FT DOMAIN 67 266  
 FT REPEAT 67 88  
 FT REPEAT 89 110  
 FT REPEAT 111 121  
 FT REPEAT 122 143  
 FT REPEAT 144 165  
 FT REPEAT 166 187

FT	REPEAT	188	209	7.
FT	REPEAT	210	231	8.
FT	REPEAT	232	242	9 (HALF-LENGTH).
FT	REPEAT	243	266	10.
FT	CONFLICT	18	18	A -> R (IN REF. 2; CAA29858).
FT	CONFLICT	44	44	MISSING (IN REF. 3).
FT	CONFLICT	45	45	V -> I (IN REF. 2; CAA29858).
FT	CONFLICT	107	107	E -> Q (IN REF. 3).
FT	CONFLICT	123	123	Y -> F (IN REF. 2; CAA29857).
FT	CONFLICT	147	147	A -> V (IN REF. 2; CAA29858 AND REF. 3).
FT	CONFLICT	150	150	R -> G (IN REF. 2; CAA29858 AND CAA29857).
FT	CONFLICT	191	191	N -> Q (IN REF. 3).
FT	CONFLICT	195	195	MISSING (IN REF. 2; CAA29858 AND CAA29857).
FT	CONFLICT	211	211	S -> K (IN REF. 2; CAA29858 AND CAA29857).
FT	CONFLICT	255	256	VL -> LV (IN REF. 3).
FT	CONFLICT	256	256	L -> V (IN REF. 2; CAA29858 AND CAA29857).
SO	SEQUENCE	266	AA; 30591	MM; OFF6DB386497C7D2 CRC64;

Query Match	77.3%;	Score 430;	DB 1;	Length 266;
Best Local Similarity	78.6%;	Pred. No. 4.6e-29;		
Matches	88;	Conservative	8;	Mismatches 16;
			Indels	0;
			Gaps	0;

```
QY 1 QKHLHLOEKLSPJGEMRDRARAHVDAIETHLAPYSDELROQLARLEALKENGARLAE 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 QKTELQEKLSPLAEELRDSARTHVDTLRTKLAPYSNELQORLAARLESIKEGGASTLAE 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 YHAKATEHLSTSEKAKPALEDLRQGLLPVLESFVSFLSALEEYTKKLTNTQ 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 YQAKAREHLSTSEKAKPALEDLRQGLLPVLESFVSFKASVQNVLTDEATKKLTNTQ 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT	5
APAL_BOVIN	
ID	APAL_BOVIN
STANDARD;	
PRT;	265 AA.

ne 01-APR-1990 (Rel. 14, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=30348478; pubmed=2117227;  
RA O'Huigin C., Chan L., Li W.H.;  
RT "Cloning and sequencing of bovine apolipoprotein A-I cDNA and  
RL molecular evolution of apolipoproteins A-I and B-100.",  
Mol. Biol. Evol. 7:327-339(1990).

SEQUENCE OF 19-265.  
MEDLINE=92153895; Pubmed=1739745;  
Sparrow D.A., Lee B.R., Laplaud M.P., Auboirn S., Bauchart D.,  
RA Chapman J.M., Gotto A.M. Jr., Yang C.Y., Sparrow J.T.;  
RT "Plasma lipid transport in the preinvariant calf, Bos spp: primary  
structure of bovine apolipoprotein A-I.";  
RL Biochim. Biophys. Acta 1123:145-150(1992).

RT from plasma high density lipoproteins in the pruruminant calf, Bos  
spp. " ;  
RT  
RT  
RA Laplaud M.P., Chapman J.M. ;  
RA "Characterization and amino-terminal sequence of apolipoprotein AI  
RA Audoiron S., Sparrow D.A., Beaubatie L., Bauchart D., Sparrow J.T.,  
RX MEDLINE=90147795; Pubmed=2105728;  
RP SEQUENCE OF 25-70.  
RP 197

```

RL Biochem. Biophys. Res. Commun. 166:833-839(1990).
CC
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC -----  
DR EMBL; M35870; AAA30381.1; -.  
DR PIR; A34649; A34649.  
DR HSSP; P02647; 1ODR.  
DR InterPro; IPR000074; Apolipoprotein.

DR pfam; PF01442; Apolipoprotein; 1.  
KW plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.

FT	1	18
SIGNAL	1	
PT	19	24
PROB		

FT	265	APOLIPROTEIN A-1.
CHAIN	25	10 X APPROXIMATE MAJORITY
ENVELOPE	24	PERCENT
FT	23	

FT	DOMAIN	67	265	10	A APPROXIMATE	INDEMN	REPEATS.
FT	REPEAT	67	88	1.			

FT	89	110	2.
REPEAT			
FT	111	121	3 (HALF-LENGTH).
REPEAT			

FT	122	143	4.
FT	144	165	5.
REPEAT			
REPEAT			

REPEAT	166	187	5.
REPEAT	166	187	6.
REPEAT	166	187	7.
REPEAT	166	187	8.
REPEAT	166	187	9.
REPEAT	166	187	10.
REPEAT	166	187	11.
REPEAT	166	187	12.
REPEAT	166	187	13.
REPEAT	166	187	14.
REPEAT	166	187	15.
REPEAT	166	187	16.
REPEAT	166	187	17.
REPEAT	166	187	18.
REPEAT	166	187	19.
REPEAT	166	187	20.
REPEAT	166	187	21.
REPEAT	166	187	22.
REPEAT	166	187	23.
REPEAT	166	187	24.
REPEAT	166	187	25.
REPEAT	166	187	26.
REPEAT	166	187	27.
REPEAT	166	187	28.
REPEAT	166	187	29.
REPEAT	166	187	30.
REPEAT	166	187	31.
REPEAT	166	187	32.
REPEAT	166	187	33.
REPEAT	166	187	34.
REPEAT	166	187	35.
REPEAT	166	187	36.
REPEAT	166	187	37.
REPEAT	166	187	38.
REPEAT	166	187	39.
REPEAT	166	187	40.
REPEAT	166	187	41.
REPEAT	166	187	42.
REPEAT	166	187	43.
REPEAT	166	187	44.
REPEAT	166	187	45.
REPEAT	166	187	46.
REPEAT	166	187	47.
REPEAT	166	187	48.
REPEAT	166	187	49.
REPEAT	166	187	50.
REPEAT	166	187	51.
REPEAT	166	187	52.
REPEAT	166	187	53.
REPEAT	166	187	54.
REPEAT	166	187	55.
REPEAT	166	187	56.
REPEAT	166	187	57.
REPEAT	166	187	58.
REPEAT	166	187	59.
REPEAT	166	187	60.
REPEAT	166	187	61.
REPEAT	166	187	62.
REPEAT	166	187	63.
REPEAT	166	187	64.
REPEAT	166	187	65.
REPEAT	166	187	66.
REPEAT	166	187	67.
REPEAT	166	187	68.
REPEAT	166	187	69.
REPEAT	166	187	70.
REPEAT	166	187	71.
REPEAT	166	187	72.
REPEAT	166	187	73.
REPEAT	166	187	74.
REPEAT	166	187	75.
REPEAT	166	187	76.
REPEAT	166	187	77.
REPEAT	166	187	78.
REPEAT	166	187	79.
REPEAT	166	187	80.
REPEAT	166	187	81.
REPEAT	166	187	82.
REPEAT	166	187	83.
REPEAT	166	187	84.
REPEAT	166	187	85.
REPEAT	166	187	86.
REPEAT	166	187	87.
REPEAT	166	187	88.
REPEAT	166	187	89.
REPEAT	166	187	90.
REPEAT	166	187	91.
REPEAT	166	187	92.
REPEAT	166	187	93.
REPEAT	166	187	94.
REPEAT	166	187	95.
REPEAT	166	187	96.
REPEAT	166	187	97.
REPEAT	166	187	98.
REPEAT	166	187	99.
REPEAT	166	187	100.

FT	188	209	7.
REPEAT			
FT	210	230	8.
REPEAT			

FT	REPEAT	231	241	9 (HALF-LENGTH).
FT	REPEAT	242	265	10.

FT	CONFLICT	185	186	QL -> HV (IN REEL
CO	SEQUENCE	265 AA.	30276 MW.	06A2681FA2ABA5

[illegible]

Query Match	73.7%;	Score 409.5;	DP
-------------	--------	--------------	----

Best Local Similarity 74.1%; Pred. No. 2.3e-  
Matches 83; Conservative 12; Mismatches

OV 1 OKI HET O E K T S P I G F E M R D R A H V D A L R T H I A P Y S D E L L

[illegible]

DB 133 QKAVDELLDNLSPFMAQZELLNKRNKANHVEIUNQZLAFISDDLE

QY 61 YHAKATEHLSTSEKAKPALEDLRQGLPVLSEFKVSFL

Db 214 YHAKASEQLKALGEKAKPVELEDLRÖGLLPVLESLSKVSIL

PECTE

```

RESULI      0
APAL_PIG

```

AC	P18648;	PRT;	265 AA.
ID	APAL_PIG	STANDARD;	

DT	01-NOV-1990	(Rel. 16, Created)
DT	01-FEB-1994	(Rel. 28, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update  
 DE annotation A-T precursor (Ano-A1)

DE APOlipoprotein A I precursor (apo A1);  
GN APOA1.

05 Sus scrofa (Pig).  
06  
07  
08  
09  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840

CC Mammalia; Eutheria; Cetartiodactyla; Suina;  
 OX NCBI TaxID=9823;

RN [1] SEQUENCE FROM N A  
 PD

RX MEDLINE=93224154; Pubmed=8468059;

RT "Characterization of the apolipoprotein AI a  
RA Birschbauer A., Knipping G., Juritsch B., Assc  
RA Birschbauer A., Knipping G., Juritsch B., Assc

100

```

RT domestic pig." ;
RL Genomics 15:643-652(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Brain;
RX MEDLINE=94125128; PubMed=8294940;
RA Moeckel B., Zinke H., Flach R., Weiss B., Weller-Guettler H.,
RT Gassen H.;
RT "Expression of apolipoprotein A-I in porcine brain endothelium in
RT vitro." ;
RL J. Neurochem. 62:788-798(1994).
RN [3]
RP SEQUENCE OF 34-265 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90132667; PubMed=2105375;
RA Weller-Guettler H., Sommerfeldt M., Papandrikopoulou A., Mischek U.,
RT Bonitz D., Frey A., Grube M., Scherer J., Gassen H.G.;
RT "Synthesis of apolipoprotein A-1 in pig brain microvascular
RT endothelial cells." ;
RL J. Neurochem. 54:444-450(1990).
RN [4]
RP SEQUENCE OF 105-265 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93154581; PubMed=8428656;
RA Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;
RT "Sequences and expression of the porcine apolipoprotein A-I and C-III
RT mRNAs." ;
RL Gene 123:173-179(1993).
RN [5]
RP SEQUENCE OF 25-265.
RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,
RA Jackson K., Gustavsson I., Rapacz J.;
RL Submitted (OCT-1995) to the SWISS-PROT data bank.
RN [6]
RP SEQUENCE OF 25-34.
RX MEDLINE=76184721; PubMed=178359;
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
RT "Characterization of the plasma lipoproteins and apoproteins of the
RT Erythrocebus patas monkey." ;
RL Biochemistry 15:1928-1933(1976).
CC -|- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE
CC LIVER.
CC -|- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L00626; AAA30992.1; -.
CC EMBL: X69477; CAA49234.1; -.
CC EMBL: X17057; -; NOT_ANNOTATED_CDS.
CC EMBL: X59414; CAA42050.1; -.
CC PIR: J00704; J00704.
CC PIR: A05311; A05311.
CC PIR: A46018; A46018.
CC PIR: S21830; S21830.
CC PIR: S31394; S31394.
CC HSSP: P02647; 10DR.
CC InterPro: IPR000074; Apolipoprotein.
CC Pfam: PF01442; Apolipoprotein; 1.
CC Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; signal.
CC SIGNAL 1 18 BY SIMILARITY.
CC PROPEP 19 24

```

FT	CHAIN	25	265	APOLIPOPROTEIN A-I.
FT	DOMAIN	67	265	10 X APPROXIMATE TANDEM REPEATS.
FT	REPEAT	67	88	1.
FT	REPEAT	89	110	2.
FT	REPEAT	111	121	3 (HALF-LENGTH).
FT	REPEAT	122	142	4.
FT	REPEAT	144	165	5.
FT	REPEAT	166	187	6.
FT	REPEAT	188	209	7.
FT	REPEAT	210	230	8.
FT	REPEAT	231	241	9 (HALF-LENGTH).
FT	REPEAT	242	265	10.
FT	CONFLICT	108	108	E -> K (IN REF. 3).
FT	CONFLICT	143	143	MISSING (IN REF. 2 AND 3).
FT	CONFLICT	173	173	D -> S (IN REF. 4).
FT	CONFLICT	180	180	E -> A (IN REF. 2 AND 3).
FT	CONFLICT	185	186	HV -> QL (IN REF. 1 AND 5).
FT	CONFLICT	209	209	G -> D (IN REF. 2 AND 3).
FT	CONFLICT	224	224	A -> G (IN REF. 4).
SQ	SEQUENCE	265 AA;	30325 MW;	2C6E578318ECF69C CRC64;

Query Match	73.5%;	Score 408.5;	DB 1;	Length 265;
Best Local Similarity	74.1%;	Pred. No. 2.8e-27;		
Matches	83;	Conservative	12;	Mismatches 16;
			Indels	1;
			Gaps	1;

QY	1	QKHELQEKLSPIGEMDRDRAHVDALFPHLAPYSDELQRRLARLEALKENGARLAE	60
		:         :            :               :	
Db	155	QKVQELQEKLSPLAEELRDRLRAHVEALRQHVAPYSDDLQRQMARFEALKEGGGS-LAE	213
QY	61	YHAKATEHLSTSEKAKPALEDLRQGLLPVLESFVSLALEEYTKKLTQ	112
		:   :   :	
Db	214	YQAKAQEQKLKALGEKAKPALEDLRQGLLPVLENLKVSITLAIIEASKKLNAQ	265

```

RESULT 7
APAL_TUPGB
ID APAL_TUPGB STANDARD; PRT; 265 AA.
AC 018759;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RL Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC THE LECTIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC THE LECTIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF005638; AAB82326.1; -.
CC HSP; P02647; IAVL.
CC InterPro; IPR000074; Apolipoprotein

```



```

DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 265 APOLIPOPROTEIN A-I.
FT DOMAIN 67 265 10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (HALF-LENGTH).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (HALF-LENGTH).
FT REPEAT 243 265 10.
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAEA165 CRC64;

Query Match 63.8%; Score 355; DB 1; Length 265;
Best Local Similarity 62.4%; Pred. NO. 7.6e-23;
Matches 68; Conservative 18; Mismatches 23; Indels 0; Gaps 0;

QY 1 OKLHELQEKLSPLGEEMRDKRAHVDALRTHLAPYSDELQKRLARLEALKENGARLAE 60
   ||| ||||::|||::||| ||: | ||| ||||::||: | ||||:| | | |||
Db 155 QKLMELQEQVTPGLGEDLRDSVRAYADTLRTQLAPYSEQMRTLGARLEAIKEGGSASLAE 214
   ||||:| | | | |||| |||: |||: | ||| |::| | |

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEYTKKL 109
   ||||:| | | | |||| |||: |||: | ||| |::| | |
Db 215 YHAKASEQLSALGEKAKPVLEDIHQGLMPMWESEFKTGVLNVIDEAKKL 263

RESULT 8
APAL_MOUSE STANDARD; PRT; 264 AA.
ID APAL_MOUSE AC 000623;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
OS APOA1.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92281682; Pubmed=1596360;
RA Stoffel W., Mueller R., Binczek E., Hofmann K.;
RT "Mouse apolipoprotein AI. cDNA-derived primary structure, gene
RT organisation and complete nucleotide sequence.";
RL Biol. Chem. Hoppe-Seyler 373:187-193(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93122774; Pubmed=1478650;
RA Januzzi J.L., Azrolan N., O'Connell A., Aalto-Setälä K., Breslow J.L.;
RT "Characterization of the mouse apolipoprotein ApoA-1/Apoc-3 gene
RT locus: genomic, mRNA, and protein sequences with comparisons to other
RT species.";
RL Genomics 14:1081-1088(1992).
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC		or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC		-----
DR	EMBL; X64262; CAA45560.1; -;	
DR	EMBL; X64263; CAA45561.1; -;	
DR	EMBL; L04149; -; NOT_ANNOTATED_CDS.	
DR	EMBL; L04151; -; NOT_ANNOTATED_CDS.	
DR	PIR; S22420; S22420.	
DR	PIR; A44364; A44364.	
DR	HSSP; P02647; IAVI.	
DR	SWISS-2DPAGE; Q00623; MOUSE.	
DR	MGD; MGI:88049; Apoal.	
DR	InterPro; IPR000074; Apolipoprotein.	
DR	Pfam; PF01442; Apolipoprotein; 1.	
KW	plasma; lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.	
FT	SIGNAL	1 18 BY SIMILARITY.
FT	PROPEP	19 24 BY SIMILARITY.
FT	CHAIN	25 264 APOLIPOPROTEIN A-I.
FT	DOMAIN	67 264 10 X APPROXIMATE TANDEM REPEATS.
FT	REPEAT	67 88 1.
FT	REPEAT	89 110 2.
FT	REPEAT	111 121 3 (HALF-LENGTH).
FT	REPEAT	122 143 4.
FT	REPEAT	144 165 5.
FT	REPEAT	166 187 6.
FT	REPEAT	188 207 7 (INCOMPLETE).
FT	REPEAT	208 229 8.
FT	REPEAT	230 240 9 (HALF-LENGTH).
FT	REPEAT	241 264 10.
SO	SEQUENCE	264 AA; 30587 MW; C453FF2019634AAC CRC64;

```

Query Match          53.2%; Score 296; DB 1; Length 264;
Best Local Similarity 53.6%; Pred. No. 5.9e-18;
Matches 60; Conservative 18; Mismatches 32; Indels 2; Gaps 1;

QY 1 OKLHELQEKLSPLGSEMRDRAAHVDAIRTHLAPYSDELQRLARLEALKENGARLAE 60
   ||| ||| :||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 QKLQELQGRLSPLVAEEFRDRMRTHVDSLRTQLAPHSEQMRSLAQRLAELKSN--PTLNE 212
   || :| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 YHAKATEHLSTLSEKAKPALEDLRLQGLLPVLESFKVSFLSALEEYTKKLTNQ 112
   || :| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 YHTRAKTHLKTGEEKARPALEDLRHSLMPLETLTKTKAQSVIDKASETLTQAQ 264
   || :| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
APAL_ANAPL
ID APAL_ANAPL STANDARD; PRT; 264 AA.
AC 042296;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEKIN BREED; TISSUE=Liver;
RA Chen B., Lu X., Wang K., Xue H., Zeng W., Zhu D., Chi L.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----

```



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88049703; PubMed=3118875;  
 RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.;  
 RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and  
 evolution.";  
 RL Biochem. Biophys. Res. Commun. 148:485-492(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88152500; PubMed=3126099;  
 RA Ferrari S., Tarugi P., Drusiani E., Calandra S., Fregni M.;  
 RT "The complete sequence of chick apolipoprotein AI mRNA and its  
 expression in the developing chick.";  
 RL Gene 60:39-46(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87222301; PubMed=3108248;  
 RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,  
 RA Leberer H., Lusis A.J.;  
 RT "Structure, evolution, and regulation of chicken apolipoprotein A-I.";  
 RL J. Biol. Chem. 262:7058-7065(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92381402; PubMed=1512510;  
 RA Lamou-Fava S., Sasstry R., Ferrari S., Rajavashisth T.B.,  
 RA Lusis A.J., Karathanasis S.K.;  
 RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene  
 expression: differences between avian and mammalian apoA-I gene  
 transcription control regions.";  
 RL J. Lipid Res. 33:831-842(1992).  
 RN [5]  
 RP SEQUENCE OF 25-44.  
 RX MEDLINE=83213468; PubMed=6406496;  
 RA Shackelford J.E., Leberer H.G.;  
 RT "Synthesis and secretion of apolipoprotein A1 by chick breast  
 muscle.";  
 RL J. Biol. Chem. 258:7175-7180(1983).  
 CC -|- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC -|- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M17961; AAA48593.1; -;  
 DR EMBL; M18746; AAA48594.1; -;  
 DR EMBL; M25559; AAA48592.1; -;  
 DR EMBL; M96012; AAA48597.1; -;  
 DR PIR; S01453; LPPCHAL.  
 DR PIR; JH0471; JH0471.  
 DR HSSP; P02647; IODR.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 1.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 264  
 FT DOMAIN 67 264  
 FT REPEAT 67 88 10 X APPROXIMATE TANDEM REPEATS.

FT REPEAT 89 110 2.  
 FT REPEAT 111 121 3 (HALF-LENGTH).  
 FT REPEAT 122 143 4.  
 FT REPEAT 144 165 5.  
 FT REPEAT 166 187 6.  
 FT REPEAT 188 209 7.  
 FT REPEAT 210 231 8.  
 FT REPEAT 232 242 9 (HALF-LENGTH).  
 FT REPEAT 243 264 10.  
 FT CONFLICT 16 16 T -> I (IN REF. 4).  
 FT CONFLICT 148 148 E -> K (IN REF. 3).  
 SQ SEQUENCE 264 AA; 30680 MW; 968320E81E2AC5C2 CRC64;  
 QY Query Match 45.9%; Score 255; DB 1; Length 264;  
 Db Best Local Similarity 46.8%; Pred. No. 1.5e-14;  
 Matches 51; Conservative 21; Mismatches 37; Indels 0; Gaps 0;  
 QY 1 OKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRRLAARLEAKENGARLAE 60  
 Db 155 QKVELMQAKLTPEAEARDRLRGHEELRKNAAPYSDELQRRLAARLEAKENGARLAE 214  
 QY 61 YHAKATEHLSTSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKL 109  
 Db 215 YQAKVMEQLSNLRKMTPLVQEFRRRLTPYAENLKNRLISFLDELQKSV 263  
 RESULT 12  
 ID APAL\_COTJA STANDARD; PRT; 264 AA.  
 AC P32918;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 OX NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=97212027; PubMed=9058967;  
 RA Oku H., Toda T., Nagata J., Ishikawa M., Neyazaki K., Shinjo C.,  
 RA Chinen I.;  
 RT "Apolipoprotein A-1 of Japanese quail: cDNA sequence and modulation of  
 RT tissue expression by cholesterol feeding.";  
 RL Biosci. Biotechnol. Biochem. 61:286-290(1997).  
 RN [2]  
 RP SEQUENCE OF 25-60.  
 RX MEDLINE=93213845; PubMed=8461329;  
 RA Oku H., Ishikawa M., Nagata J., Toda T., Chinen I.;  
 RT "Lipoprotein and apoprotein profile of Japanese quail.";  
 RL Biochim. Biophys. Acta 1167:22-28(1993).  
 CC -|- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- TISSUE SPECIFICITY: MAJOR PROTEIN OF VLDL, HDL, LDL AND IN  
 CC CHYLOMICRONS. EXPRESSED IN A NUMBER OF TISSUES INCLUDING LIVER,  
 CC SMALL INTESTINE, LUNG, KIDNEY, HEART AND MUSCLE WITH HIGHEST  
 CC EXPRESSION IN LIVER AND SMALL INTESTINE.  
 CC -|- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>





RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
RA Ferrell R.E., Pollitzer W.S.;  
RT "Molecular basis of a unique African variant (A-IV 5) of human  
RT apolipoprotein A-IV and its significance in lipid metabolism.";  
RL Genet. Epidemiol. 9:379-388(1992).  
RN [15]  
RP VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
RX MEDLINE=95245341; PubMed=7728150;  
RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
RA Csaszar A.;  
RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
RT frequencies, effect on lipid levels, and sequence of two new  
RT variants.";  
RL Hum. Mutat. 5:58-65(1995).  
RN [16]  
RP VARIANTS FCHL SEATTLE SER-161; LEU-178 AND GLN-264.  
RX MEDLINE=97114287; PubMed=8956036;  
RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;  
RT "Two novel apolipoprotein A-IV variants in individuals with familial  
RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
RT activity.";  
RL Hum. Mutat. 8:319-325(1996).  
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
CC COMPONENT OF HDL AND CHYLOMICRONS.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
CC SECRETED IN PLASMA.  
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC -1- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV\*0  
CC TO APOA-IV\*7). APOA-IV\*1 IS THE MAJOR ALLELE (90%), IV\*2 IS ALSO  
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.  
CC -1- DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN  
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF  
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M14642; AAA51745.1; -  
DR EMBL; X13629; CAA31955.1; -  
DR EMBL; M14566; AAA51748.1; -  
DR EMBL; J02758; AAA96731.1; -  
DR EMBL; M13654; AAA51744.1; -  
DR PIR; A26481; LPHU4.  
DR PIR; A24449; A24449.  
DR PIR; A29330; A29330.  
DR PIR; A26280; A26280.  
DR PIR; S02715; S02715.  
DR HSSP; P02649; INFO.  
DR SWISS-2DPAGE; P06727; HUMAN.  
DR MIM; 107690; -  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;  
KW Disease mutation; Polymorphism.  
FT SIGNAL 1 20  
FT CHAIN 21 396 APOLIPOPROTEIN A-IV.  
FT DOMAIN 33 330 13 X 22 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 33 54 1.  
FT REPEAT 60 81 2.

FT REPEAT 82 103 3.  
FT REPEAT 115 136 4.  
FT REPEAT 137 158 5.  
FT REPEAT 159 180 6.  
FT REPEAT 181 202 7.  
FT REPEAT 203 224 8.  
FT REPEAT 225 246 9.  
FT REPEAT 247 268 10.  
FT REPEAT 269 286 11.  
FT REPEAT 287 308 12.  
FT REPEAT 309 330 13.  
FT DOMAIN 372 389 GLU/GLN-RICH.  
FT VARIANT 13 V -> M (IN APOA-IV\*1D).  
FT VARIANT 44 /FTID=VAR\_000626.  
FT VARIANT 44 E -> K (IN BUDAPEST-2).  
FT VARIANT 147 /FTID=VAR\_000627.  
FT VARIANT 147 N -> S (IN APOA-IV\*1B).  
FT VARIANT 161 /FTID=VAR\_000628.  
FT VARIANT 161 A -> S (IN SEATTLE-3; IN FCHL).  
FT /FTID=VAR\_000629.

Query Match 22.3%; Score 124; DB 1; Length 396;  
Best Local Similarity 28.6%; Pred. No. 0.0017;  
Matches 30; Conservative 22; Mismatches 53; Indels 0; Gaps 0;

QY 5 ELQEKLSPLGEEMDRRAHVDAALRTHLAPYSDELQRRLARLALKENGARLAHYHAK 64  
Db 53 ELTQQLNALFQDKLGEVNTYAGDLQKKLVPRATELHERLAKDSEKLEIGKELEELRAR 112  
QY 65 ATEHLSTSEKAKPALEDLRQGLLPVLESEKVSFLSALEYTKKL 109  
Db 113 LHPHANEVSQKIGDNLRELQORLEPYADQLRTQVNTQAEQLRRQL 157

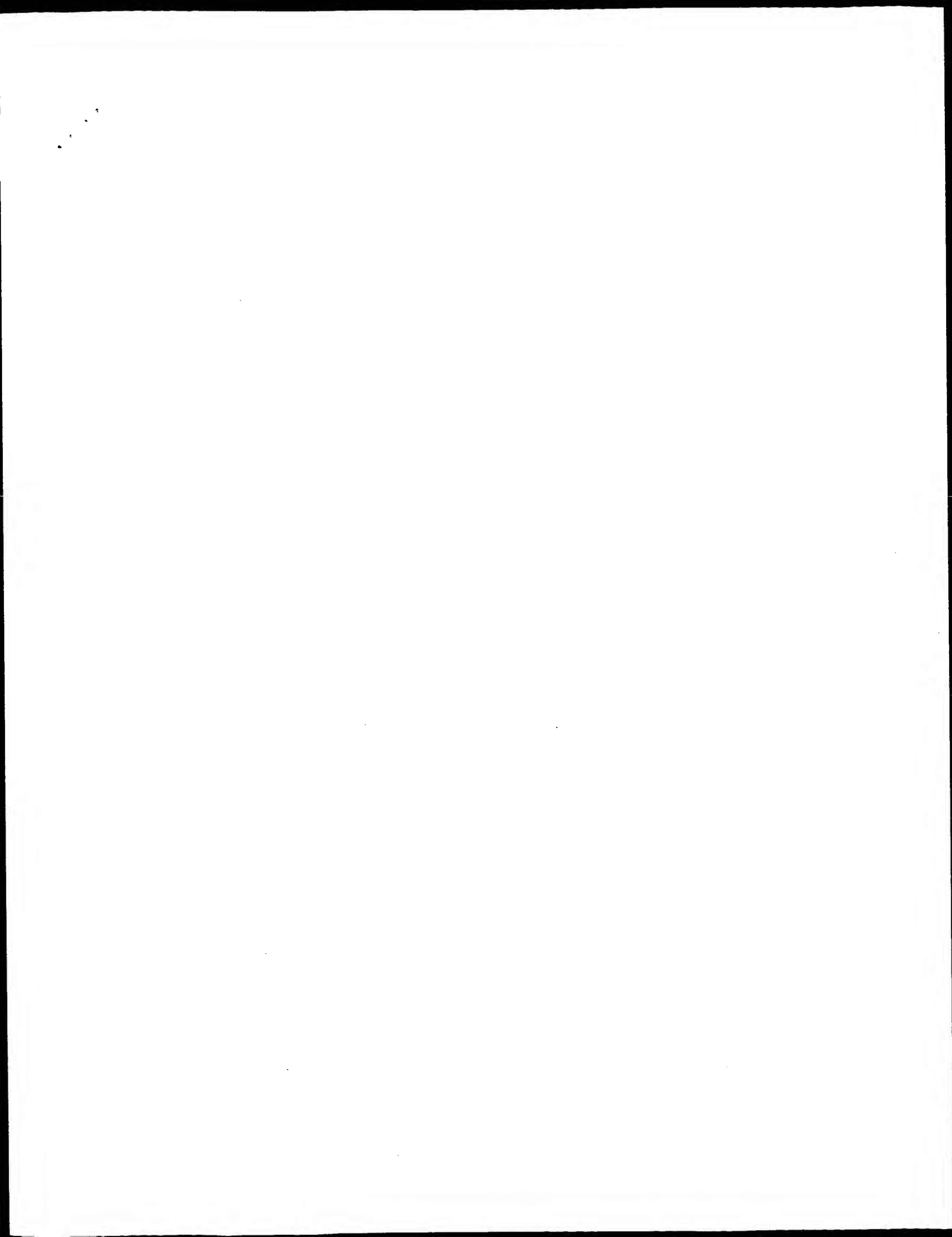
RESULT 14  
APAL\_BRARE  
ID APAL\_BRARE STANDARD; PRT; 262 AA.  
AC 042363;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=97385153; PubMed=9238027;  
RA Babin P.J., Thisse C., Durliat M., Andre M., Akimenko M.-A.,  
RA Thisse B.;  
RT "Both apolipoprotein E and A-I genes are present in a nonmammalian  
RT vertebrate and are highly expressed during embryonic development.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:8622-8627(1997).  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT) (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN THE YOLK SYNCTIAL LAYER  
CC DURING EMBRYONIC (STARTING AT THE GASTRULA STAGE) AND EARLY LARVAL  
CC DEVELOPMENT, AN EXTRAEMBRYONIC STRUCTURE IMPLICATED IN EMBRYONIC  
CC AND LARVAL NUTRITION.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial



Sun Sep 22 12:10:13 2002

us-09-803-918a-2\_copy\_156\_267.rsp

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:21:37 ; Search time 111.17 Seconds  
(without alignments)  
174.287 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_156\_267  
Perfect score: 556  
Sequence: 1 OKLHELQEKLSPLGEMRDR.....SEKVSFLSALFEYTKKLTNTQ 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	367	66.0	264	11	Q9Z2L4 mesocricetu
2	327	58.8	67	4	Q9Y355 homo sapien
3	297	53.4	263	11	008855 mus musculu
4	296	53.2	263	11	009042 mus musculu
5	287.5	51.7	241	6	Q9TS49 erinaceus e
6	273	49.1	258	11	009054 ratu
7	271	48.7	258	11	008877 ratu
8	132	23.7	263	13	Q98TG6 anguilla ja
9	130.5	23.5	366	13	Q93601 gallus gall
10	113	20.3	174	13	Q98S13 cyprinus ca
11	112.5	20.2	244	4	Q13784 homo sapien
12	111	20.0	1547	5	Q26471 schistocerc
13	109.5	19.7	102	13	Q98TJ9 platichthys
14	102.5	18.4	395	11	Q9DBN0 mus musculu
15	102.5	18.4	395	11	Q91XF8 mus musculu
16	102.5	18.4	435	11	Q01488 mus musculu

17	97.5	17.5	224	13	Q9PT73 scophthalmu
18	97	17.4	317	6	Q9GLM8 gorilla gor
19	95	17.1	313	6	Q9GLC0 tupia glis
20	93	16.7	317	6	Q9GJ03 pan troglod
21	93	16.7	317	6	Q9GLM7 pongo pygma
22	93	16.7	317	6	Q9GLM6 hylobates l
23	91.5	16.5	113	3	Q08245 saccharomyc
24	90.5	16.3	1518	5	Q21442 caenorhabdi
25	90.5	16.3	3704	5	P91904 caenorhabdi
26	89.5	16.1	1845	12	Q9WJ28 human cytom
27	88	15.8	363	4	Q9UBJ3
28	87.5	15.7	194	11	Q9D2P9
29	87	15.6	229	16	Q9CNI8
30	87	15.6	2310	5	Q9NKR1
31	87	15.6	2354	5	Q9NKR9
32	86.5	15.6	240	16	Q9RXL9
33	86	15.5	457	11	Q9DCV7
34	86	15.5	731	5	Q9N8D2
35	85.5	15.4	585	4	Q9UH65
36	85.5	15.4	585	4	Q9P0Z8
37	85.5	15.4	603	4	Q75135
38	84.5	15.2	500	5	Q9GT16
39	84.5	15.2	585	11	Q08443
40	84	15.1	259	13	Q98TG5
41	84	15.1	1175	5	Q9TY21
42	84	15.1	1201	5	Q18392
43	84	15.1	1411	4	Q15075
44	84	15.1	2067	5	Q9VJ13
45	83.5	15.0	1164	5	P92021

ALIGNMENTS

RESULT	ID	Q9Z2L4	PRELIMINARY;	PRT;	264 AA.
AC	Q9Z2L4	01-MAY-1999 (TREMBlrel. 10, Created)			
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)				
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)				
DE	APOLIPROTEIN A-I.				
GN	APOAI.				
OS	Mesocricetus auratus (Golden hamster).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;				
OC	Mesocricetus.				
OX	NCBI_TaxID=10036;				
RN	[1]	SEQUENCE FROM N.A.			
RP	STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;				
RC	MEDLINE=99061559; PubMed=9843713;				
RX	Wu J.Y.J., Reaves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;				
RA	"Zinc deficiency decreases plasma level and hepatic mRNA abundance of				
RT	apolipoprotein A-I in rats and hamsters."				
RL	Am. J. Physiol. 275:C1516-C1525(1998).				
DR	EMBL; AF046919; AAC98484.1; -				
DR	HSSP; P02647; IAV1.				
DR	InterPro; IPR000074; Apolipoprotein.				
DR	Pfam; PF01442; Apolipoprotein; 1.				
KW	lipoprotein.				
SQ	SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;				

Query Match 66.0%; Score 367; DB 11; Length 264;  
Best Local Similarity 65.2%; Pred. No. 4e-23;  
Matches 73; Conservative 13; Mismatches 24; Indels 2; Gaps 1;

Oy	1	OKLHELQEKLSPLGEMRDRARAHDALRTHLAPYSDELQRRLAARLEALKEGARGARLAE 60
Db	155	OKLHELQEKLTPLGEDLRMRHVDALRTKMTPYSDQMRDLAERLAQLKDS--PTLAE 212
Oy	61	YHAKTEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALFEYTKKLTNTQ 112

Db 213 YHTKADHLKAFGEKAKPALEDLRQGLMPVFESFKTRIMSMVEEASKKLNAQ 264

RESULT 2  
ID 09Y355 PRELIMINARY; PRT; 67 AA.  
AC 09Y355;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE APOLIPROTEIN A1 (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99216322; PubMed=10198255;  
RA Hamidi Asl K., Ilepnieks J.J., Nakamura M., Parker F., Benson M.D.;  
RT "A novel apolipoprotein A-1 variant, Arg173Pro, associated with  
cardiac and cutaneous amyloidosis."  
RL Biochem. Biophys. Res. Commun. 257:584-588(1999).  
DR EMBL; AF148963; AAD34604.1; -.  
DR HSSP; P02647; 1AV1.  
KW Lipoprotein.  
FT NON\_TER 1 1  
SQ SEQUENCE 67 AA; 7433 MW; 525E1FEB7BDD5AFB CRC64;

Query Match 58.8%; Score 327; DB 4; Length 67;  
Best Local Similarity 98.5%; Pred. No. 1.8e-20;  
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 LSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLA EYHAKATEHL 69  
Db 1 LSPLGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLA EYHAKATEHL 60

QY 70 STLSEKA 76  
Db 61 STLSEKA 67

RESULT 3  
ID 008855 PRELIMINARY; PRT; 263 AA.  
AC 008855;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6; TISSUE=SPLEEN;  
MEDLINE=98077648; PubMed=9415807;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
gene."  
RL Biochem. Mol. Biol. Int. 43:989-996(1997).  
DR EMBL; U79574; AAB58426.1; -.  
DR HSSP; P02647; 1AV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 263 AA; 30544 MW; 7825DF7483A5B24A CRC64;

Query Match 53.4%; Score 297; DB 11; Length 263;

Best Local Similarity 53.6%; Pred. No. 2.7e-17;  
Matches 60; Conservative 18; Mismatches 32; Indels 2; Gaps 1;

QY 1 OKLHELQEKLSPLGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLA E 60  
Db 154 OKLQELQRLSPVAEEFRDMRTHVDSLRTQLAPHSEQMRSLAQRLAEIKSN--PTLNE 211  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLPVLESFKVSFLSAL EYTKKLTQ 112  
Db 212 YHTRAKTHLKTGKARPALEDLRHSLMPLETLTKTQVSVIDKASETLTAQ 263

RESULT 4  
ID 009042 PRELIMINARY; PRT; 263 AA.  
AC 009042;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ICR AND BALB/C, AND C3H; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.  
DR EMBL; U79575; AAB58427.1; -.  
DR EMBL; U79572; AAB58424.1; -.  
DR EMBL; U79573; AAB58425.1; -.  
DR HSSP; P02647; 1AV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 263 AA; 30516 MW; E245DF7483A5A0DD CRC64;

Query Match 53.2%; Score 296; DB 11; Length 263;  
Best Local Similarity 53.6%; Pred. No. 3.3e-17;  
Matches 60; Conservative 18; Mismatches 32; Indels 2; Gaps 1;

QY 1 OKLHELQEKLSPLGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLA E 60  
Db 154 OKLQELQRLSPVAEEFRDMRTHVDSLRTQLAPHSEQMRSLAQRLAEIKSN--PTLNE 211  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLPVLESFKVSFLSAL EYTKKLTQ 112  
Db 212 YHTRAKTHLKTGKARPALEDLRHSLMPLETLTKTQVSVIDKASETLTAQ 263

RESULT 5  
ID 09TS49 PRELIMINARY; PRT; 241 AA.  
AC 09TS49;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I, APOA-I=CHOLESTEROL TRANSPORTER.  
OS Erinaceus europaeus (Western European hedgehog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceus.  
OX NCBI\_Taxid=9365;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95294458; PubMed=7775860;  
RA Sparrow D.A., Laplaud P.M., Saboureau M., Zhou G., Dolphin P.J.,  
Gotto A.M., Jr., Sparrow J.T.;  
RT "Plasma lipid transport in the hedgehog: partial characterization of  
structure and function of apolipoprotein A-I."  
RL J. Lipid Res. 36:485-495(1995).  
DR HSSP; P02647; 1AV1.

DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
SQ SEQUENCE 241 AA; 27630 MW; 2EF00F2B69210535 CRC64;

Query Match 51.7%; Score 287.5; DB 6; Length 241;  
Best Local Similarity 53.6%; Pred. No. 1.5e-16;  
Matches 60; Conservative 14; Mismatches 37; Indels 1; Gaps 1;

QY 1 OKLHELQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGARLAE 60  
DB 131 OKAQLQOKAGELGQOHDRVRTHVDALRTDLAPYGGEARKLLQRLQDIKAKSG-DLAE 189

QY 61 YHAKATEHLSLSEKAKPALEDLRQGLLPVLESFVSLSALEYTKKLNQ 112  
DB 190 YQTKLSEHLKSPFEKAPTLQDLRHGLEPLWEGTKAGAMSMLEELGKKLNSQ 241

RESULT 6  
ID 009054 PRELIMINARY; PRT; 258 AA.  
AC 009054;

DT 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE APOLIPROTEIN A-I.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WKY, AND SHRSF; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79578; AAB58430.1; -  
DR EMBL; U79577; AAB58429.1; -  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1. \*  
KW Lipoprotein.  
SQ SEQUENCE 258 AA; 29918 MW; 093E6EF2E629CDC8 CRC64;

Query Match 49.1%; Score 273; DB 11; Length 258;  
Best Local Similarity 50.4%; Pred. No. 2.6e-15;  
Matches 57; Conservative 18; Mismatches 32; Indels 6; Gaps 2;

QY 2 KLH---ELQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGAR 57  
DB 147 ELHKNAKEMQRLKLVAAEFDRMRVNDALRAKFGLYSDQMRNLAQRLTEIKNH--PT 204

QY 58 LAEYHAKATEHLSLSEKAKPALEDLRQGLLPVLESFVSLSALEYTKKLN 110  
DB 205 LIEYHTKASDHLKLTGKAKAPALDLGGLMPVLEAWKAKIMSMIDEAKKKLN 257

RESULT 7  
ID 008877 PRELIMINARY; PRT; 258 AA.  
AC 008877;  
DT 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SHR; TISSUE=SPLEEN;  
RX MEDLINE=98077648; PubMed=9415807;

RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
RT gene.";  
RL Biochem. Mol. Biol. Int. 43:989-996(1997).

DR EMBL; U79576; AAB58428.1; -  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.

SQ SEQUENCE 258 AA; 29831 MW; 093FB582E629CDC8 CRC64;

Query Match 48.7%; Score 271; DB 11; Length 258;  
Best Local Similarity 50.4%; Pred. No. 3.9e-15;  
Matches 57; Conservative 17; Mismatches 33; Indels 6; Gaps 2;

QY 2 KLH---ELQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGAR 57  
DB 147 ELHKNAKEMQRLKLVAAEFDRMRVNDALRAKFGLYSDQMRNLAQRLTEIKNH--PT 204

QY 58 LAEYHAKATEHLSLSEKAKPALEDLRQGLLPVLESFVSLSALEYTKKLN 110  
DB 205 LIEYHTKASDHLKLTGKAKAPALDLGGLMPVLEAWKAKIMSMIDEAKKKLN 257

RESULT 8  
ID 098TG6 PRELIMINARY; PRT; 263 AA.  
AC 098TG6;

DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)

DE 28KDA-2 APOLIPROTEIN.  
OS Anguilla japonica (Japanese eel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
OC Anguillidae; Anguilla.  
OX NCBI\_TaxID=7937;

RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=21175592; PubMed=11278178;  
RA Kondo H., Kawazoe I., Nakaya M., Kikuchi K., Aida K., Watabe S.;  
RT "The novel sequences of major plasma apolipoproteins in the eel  
RT Anguilla japonica.";  
RL Biochim. Biophys. Acta 1531:132-142(2001).  
DR EMBL; AB046203; BAB40960.1; -  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 263 AA; 30087 MW; DB8CB0AF30EBE6B1 CRC64;

Query Match 23.7%; Score 132; DB 13; Length 263;  
Best Local Similarity 29.0%; Pred. No. 0.0015;  
Matches 31; Conservative 26; Mismatches 48; Indels 2; Gaps 1;

QY 1 OKLHELQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRRLAARLEALKENGARLAE 60  
DB 155 QEMELKTKLQVYEDLRARIQVNVETKSKLVPIVEAIRAKLTERLEELRTLAEPYOE 214

QY 61 YHAKATEHLSLSEKAKPALEDLRQGLLPVLESFVSLSALEYTK 107  
DB 215 YKDHLSEALTDVKDVQG--EDLQSKIKPYAEELKTKLVALMESLSQ 259

RESULT 9  
ID 093601 PRELIMINARY; PRT; 366 AA.  
AC 093601;

DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE APOLOPROTEIN AIV.  
GN APOAIV.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98221191; PubMed=9553114;  
RA Steinmetz A., Hermann M., Nimpf J., Aebersold R., Ducret A.,  
RT Weinberg R.B., Schneider W.J.;  
RT "Expression and conservation of apolipoprotein AIV in an avian  
RT species";  
RL J. Biol. Chem. 273:10543-10549(1998).  
DR EMBL: Y16534; CAA76273.1; -.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 366 AA; 40852 MW; B47DB49F3E8D91CE CRC64;

Query Match 23.5%; Score 130.5; DB 13; Length 366;  
Best Local Similarity 25.9%; Pred. No. 0.0029;  
Matches 36; Conservative 27; Mismatches 43; Indels 33; Gaps 2;

QY 6 LQEKLSPLGEEMDRARAHVDALRTHLAPYSDELRL-----QLAARLEALKENGARL 58  
DB 196 LKQGLTPLADELKEQVAQSEGLRKGLSPYAQEVQDGLNRQLSLTAQMERAAEELRSRL 255  
QY 59 AEYHAKATEHLSTLSEKAKPAL-----EDLRQGLLPVLE 92  
DB 256 AASSEEMRAQLSPLAQLQELRGDAEAMQRLAPLAQGLDERLAQTVEAFRQAAPISE 315  
QY 93 SEKVSFLSALLEEYTKKLT 111  
DB 316 TFRQQLVORLEEMKOKLES 334

## RESULT 10

Q98S13 PRELIMINARY; PRT; 174 AA.  
ID Q98S13;  
AC Q98S13;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE APOLOPROTEIN A-I (FRAGMENT).  
GN APOA1.  
OS Cyprinus carpio (Common carp).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Cyprinus.  
OX NCBI\_TaxID=7962;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Concha M.I., Villanueva J., Amthauer R.J.;  
RT "Synthesis and secretion of apolipoprotein A-I in the carp skin: A  
RT possible defensive mechanism";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ308993; CAC34942.1; -.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 174 AA; 20810 MW; 435FC05AE0EA0578 CRC64;

Query Match 20.3%; Score 113; DB 13; Length 174;  
Best Local Similarity 23.4%; Pred. No. 0.035;  
Matches 30; Conservative 27; Mismatches 51; Indels 20; Gaps 2;

QY 3 LHELQEKLSPLGEEMDRARAHVDALRTHLAP-----YSDELRLAARLEALK 51

DB 41 VEELRKKIEPMRAELROVLEKHLQEYRDELPVEEYLLTKHQKFLLEMRKLEPVVKSRL 100  
QY 52 ENGARLAEEYHAKATEHLSTLSEKAKPALEDRLQGLLPVLESFKVSFLSALLEY----- 105  
DB 101 EKFGPNWEETKSKLMPLEAVREKVAEHLQDLKLLPEYMQDYREQMEKGAQEFQSVKS 160  
QY 106 ---TKKLN 110  
DB 161 GELRKKMN 168

## RESULT 11

Q13784 PRELIMINARY; PRT; 244 AA.  
ID Q13784;  
AC Q13784;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE APOA4 PROTEIN (FRAGMENT).  
GN APOA4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=86016704; PubMed=3931073;  
RA Karathanasis S.K.;  
RT "Apolipoprotein multigene family: tandem organization of human  
RT apolipoprotein AI, CII, and AIV genes";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:6374-6378(1985).  
DR EMBL: M10373; AAB59516.1; -.  
DR HSSP: P02649; 1BZ4.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
FT NON\_TER 1  
SQ SEQUENCE 244 AA; 28157 MW; 01616ACB53F12734 CRC64;

Query Match 20.2%; Score 112.5; DB 4; Length 244;  
Best Local Similarity 25.4%; Pred. No. 0.057;  
Matches 35; Conservative 27; Mismatches 47; Indels 29; Gaps 3;

QY 1 QKLHELQEKLSPLGEEMDRARAHVDALRTHLAPYSDELRLQRLAARLEAL-----KENGGA 56  
DB 58 QNVEELKGRITPYADEFKKIDQYVEELRLSLAPYAQDTQEKLNHQLEGLTFQMKNAEE 117  
QY 57 RLAEYHAKATE---HLSTLSEKAKPAL-----EDLRQGLLPVL 91  
DB 118 LKARISASAEELRQRLAPLAEDVRGNLKGNTGLQKSLAELGHLDDQVEEFRRAVEBYG 177  
QY 92 SEKVSFLSALLEEYTKKL 109  
DB 178 ENFNKALVQOMEQLRQKL 195

## RESULT 12

Q26471 PRELIMINARY; PRT; 1547 AA.  
ID Q26471;  
AC Q26471;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE BLACKJACK PROTEIN.  
OS Schistocerca americana (American grasshopper).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;  
OC Acridomorpha; Acridoidea; Acrididae; Schistocerca.  
OX NCBI\_TaxID=7009;  
RN [1]  
RP SEQUENCE FROM N.A.



RC TISSUE=EMBRYO;  
RX MEDLINE=96393048; PubMed=8799836;  
RA Zachow K.R., Bentley D.;  
RT "Blackjack, a novel protein associated with microtubules in embryonic  
neurons.";  
RL J. Cell Sci. 109:1497-1507(1996).  
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.  
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, HEAVILY EXPRESSED IN CNS AND  
CC PNS AXONS. EXPRESSED AT LOWER LEVELS IN EPITHELIAL AND MESODERMAL  
CELLS.  
DR EMBL; L76606; AAC37266.1; -  
DR InterPro; IPR004155; HEAT\_PBS.  
DR Pfam; PF03130; HEAT\_PBS; 7.  
KW Coiled coil; Microtubules; Repeat.  
FT DOMAIN 630 860 21 X 11 AA TANDEM REPEATS OF P-L-E-E-L-R-  
FT REPEAT 630 640 K-D-A-A-E.  
FT REPEAT 641 651 1.  
FT REPEAT 652 662 2.  
FT REPEAT 663 673 3.  
FT REPEAT 674 684 4.  
FT REPEAT 685 695 5.  
FT REPEAT 696 706 6.  
FT REPEAT 707 717 7.  
FT REPEAT 718 728 8.  
FT REPEAT 729 739 9.  
FT REPEAT 740 750 10.  
FT REPEAT 751 761 11.  
FT REPEAT 762 772 12.  
FT REPEAT 773 783 13.  
FT REPEAT 784 794 14.  
FT REPEAT 795 805 15.  
FT REPEAT 806 816 16.  
FT REPEAT 817 827 17.  
FT REPEAT 828 838 18.  
FT REPEAT 839 849 19.  
FT REPEAT 850 860 20.  
FT DOMAIN 631 860  
SQ SEQUENCE 1547 AA; 169310 MW; 94F3D243768FDEB CRC64;  
COILED COIL (POTENTIAL).  
Query Match 20.0%; Score 111; DB 5; Length 1547;  
Best Local Similarity 31.8%; Pred. No. 0.62;  
Matches 34; Conservative 16; Mismatches 57; Indels 0; Gaps 0;  
QY 1 OKLHELOEKISPLGEEMDRARAHVDALRTHLAPYSDELRLARLEALKENGARLAE 60  
DB 640 EPLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLE 699  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEYTK 107  
DB 700 LRRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRK 746  
RESULT 13  
Q98TJ9 PRELIMINARY; PRT; 102 AA.  
Q98TJ9  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOLIPOPROTEIN A1 (FRAGMENT).  
GN APOA1.  
OS Platicthys flesus (European flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidei; Pleuronectidae; Platicthys.  
OX NCBI\_TaxID=8260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Williams T.D., Chipman J.K.;

RT "A DNA array to monitor the effects of environmental pollution on  
European flounder (Platicthys flesus).";  
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AJ310423; CAC27154.1; -  
KW Lipoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 102 AA; 11248 MW; 8F0EB2150625EDFA CRC64;

Query Match 19.7%; Score 109.5; DB 13; Length 102;  
Best Local Similarity 27.6%; Pred. No. 0.038;  
Matches 27; Conservative 20; Mismatches 40; Indels 11; Gaps 1;

QY 12 PLGEEMDRARAHVDALRTHLAPYSDELRLARLEALKENGARLAE----- 61  
DB 2 PVDAMKTSVAANVETKTALMPILLESVRAKLSERLESLEKEMATPYGKYKQQLQAYGQ 61  
QY 62 -HAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSF 98  
DB 62 AQAVKPEDLTALKKEKISPMAEVAKAVTEFMAIATF 99

RESULT 14  
Q9DBNO PRELIMINARY; PRT; 395 AA.  
Q9DBNO  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE ADULT MALE LIVER CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
DE CLONE:1300002K10, FULL INSERT SEQUENCE.  
GN APOA4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=LIVER;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nakaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzairelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse CDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK004856; BAB23620.1; -  
DR MGI; MGI:88051; APOA4.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
SQ SEQUENCE 395 AA; 45044 MW; 4102D84ACB0D182A CRC64;

Query Match 18.4%; Score 102.5; DB 11; Length 395;  
Best Local Similarity 19.8%; Pred. No. 0.67;  
Matches 23; Conservative 39; Mismatches 47; Indels 7; Gaps 1;  
QY 1 OKLHELOEKISPLGEEMDRARAHVDALRTHLAPYSDELRLARLEALKENGARLAE 60  
DB 640 EPLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLE 699  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEYTK 107  
DB 700 LRRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRKDAEPLLEELRK 746  
RESULT 13  
Q98TJ9 PRELIMINARY; PRT; 102 AA.  
Q98TJ9  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOLIPOPROTEIN A1 (FRAGMENT).  
GN APOA1.  
OS Platicthys flesus (European flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidei; Pleuronectidae; Platicthys.  
OX NCBI\_TaxID=8260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Williams T.D., Chipman J.K.;



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:05:15 ; Search time 139.03 Seconds

(without alignments)  
89.479 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_156\_267

Perfect score: 556

Sequence: 1 QKHELEKLSPLGEMRDR.....SPKVSFLSALAEYTKKLTNQ 112

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	556	100.0	243	9 AAP81082	Sequence of mature
2	556	100.0	244	22 AAU28184	Novel human secret
3	556	100.0	267	7 AAP61079	Assumed human apol
4	556	100.0	267	9 AAP82128	Entire human prepr
5	556	100.0	267	14 AAR34032	Sequence of apo AI
6	556	100.0	267	16 AAR72705	Human apo A-I incl
7	556	100.0	267	20 AAY18675	Human apolipoprote
8	556	100.0	267	22 AAB47620	Full length Apo-AI
9	556	100.0	268	9 AAP80668	Recombinant human
10	548	98.6	267	18 AAM08602	Human apolipoprote
11	545	98.0	264	15 AAR56863	Apo-lipoprotein AI

12	545	98.0	264	15 AAR56864	Apo-lipoprotein AI
13	522	93.9	299	22 AAU33170	Novel human secret
14	496.5	89.3	359	22 AAU30470	Novel human secret
15	409.5	73.7	318	22 AAU30268	Novel human secret
16	376	67.6	221	22 AAU29835	Novel human secret
17	194	34.9	151	22 AAU02278	Human polypeptide
18	187.5	33.7	119	22 AAU30468	Novel human secret
19	125	22.5	333	14 AAR39497	Human apolipoprote
20	125	22.5	342	14 AAR39498	Human apolipoprote
21	125	22.5	396	22 AAB90664	Human secreted pro
22	124	22.3	183	14 AAR39482	Human apolipoprote
23	124	22.3	333	14 AAR39481	Human apolipoprote
24	124	22.3	333	14 AAR39488	Human apolipoprote
25	124	22.3	333	14 AAR39490	Human apolipoprote
26	124	22.3	333	14 AAR39495	Human apolipoprote
27	124	22.3	337	14 AAR39492	Human apolipoprote
28	124	22.3	342	14 AAR39487	Human apolipoprote
29	124	22.3	342	14 AAR39489	Human apolipoprote
30	124	22.3	342	14 AAR39491	Human apolipoprote
31	124	22.3	342	14 AAR39496	Human apolipoprote
32	124	22.3	346	14 AAR39493	Human apolipoprote
33	124	22.3	363	14 AAR39479	Human apolipoprote
34	124	22.3	373	14 AAR39485	Human apolipoprote
35	124	22.3	377	14 AAR39443	Human apolipoprote
36	124	22.3	377	14 AAR39502	Human apolipoprote
37	124	22.3	377	14 AAR39501	Human apolipoprote
38	124	22.3	377	14 AAR45242	Human apolipoprote
39	124	22.3	377	14 AAR45243	Human apolipoprote
40	124	22.3	377	14 AAR45244	Human apolipoprote
41	122	21.9	337	14 AAR39494	Human apolipoprote
42	121	21.8	377	14 AAR39500	Human apolipoprote
43	120.5	21.7	328	14 AAR39484	Human apolipoprote
44	120.5	21.7	337	14 AAR39485	Human apolipoprote
45	120.5	21.7	363	14 AAR39478	Human apolipoprote

#### ALIGNMENTS

RESULT 1

AAAP81082

ID AAP81082 standard; protein; 243 AA.

XX AAP81082;

AC

XX

DT 14-JAN-1991 (first entry)

XX

DE Sequence of mature human apolipoprotein AI (apoAI).

XX

KW Atherosclerosis; therapy; cardiovascular disease.

XX

OS Homo sapiens.

XX

PN WO8803166-A.

XX

PD 05-MAY-1988.

XX

PF 21-OCT-1987; 87WO-EP00621.

XX

PR 23-OCT-1986; 86GB-0025435.

XX

PA (FARM ) FARMITALIA C ERBA SPA.

XX

PI Lorenzetti R, Monaco L, Soria M, Palomba R, Isacchi A, Sarmientos P;

XX

DR WPI; 1988-133240/19.

XX

DR N-PSDB; AAN80243.

XX

PT Recombinant human apo-lipoprotein AI -

PT used to lower plasma cholesterol and/or tri glyceride levels and

XX to combat atherosclerosis and cardiovascular diseases

PS Disclosure; Fig 1; 51pp; English.

```

XX The protein comprising apo AI genetic variants may be used to lower
CC plasma cholesterol and/or triglyceride levels. They may also be used to
CC combat atherosclerosis and cardiovascular diseases such as coronary
CC heart disease. Preld. proteins are Met-apo AI, Met-apo AI-T6, Met-apo
CC AI-MI and Met-apo AI-T6/MI.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 556; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 9.6e-49;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OKLHELQEKLSPLGEMRDRAHVDALRTHLAPYSDELRORLAARLEALKENGARLAE 60
Db 132 qklhelqeklsplgeemrdrarahvdalrthlapydelrqlaarlalkenggarlae 191

OY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTNQ 112
Db 192 yhakatehlstlsekakpaledlrqglpvlsefkvsflsaleeytkkltntq 243

RESULT 2
AAU28184
ID AAU28184 standard; Protein; 244 AA.
XX
AC AAU28184;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secretory protein, Seq ID No 353.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen.
XX
OS Homo sapiens.
XX
PN WO200166689-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US04942.
XX
PR 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR MPI; 2001-589934/66.
DR N-PSDB; AAS45084.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders.
XX
PS Example 5; SEQ ID No 353; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)

```

```

CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.
XX
SQ Sequence 244 AA;

Query Match 100.0%; Score 556; DB 22; Length 244;
Best Local Similarity 100.0%; Pred. No. 9.7e-49;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OKLHELQEKLSPLGEMRDRAHVDALRTHLAPYSDELRORLAARLEALKENGARLAE 60
Db 133 qklhelqeklsplgeemrdrarahvdalrthlapydelrqlaarlalkenggarlae 192

OY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTNQ 112
Db 193 yhakatehlstlsekakpaledlrqglpvlsefkvsflsaleeytkkltntq 244

RESULT 3
AAP61079
ID AAP61079 standard; Protein; 267 AA.
XX
AC AAP61079;
XX
DT 07-OCT-1991 (first entry)
XX
DE Assumed human apolipoprotein A-1 derivative gene product.
XX
KW Hyperlipaemia; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN JP61096998-A.
XX
PD 15-MAY-1986.
XX
PF 16-OCT-1984; 84JP-0216988.
XX
PR 16-OCT-1984; 84JP-0216988.
XX
PA (MITU ) MITSUBISHI CHEM IND KK.
XX

```



DR WPI; 1986-165025/26.  
DR N-PSDB; AAN60886.  
XX  
PT Human apo:lipoprotein A-1 (deriv.) prepn. - by providing DNA  
PT fragment in cloning site downstream of expression vector promoter  
PT and introducing into host microorganism.  
XX  
PS Disclosure; Fig 2; 9pp; Japanese.  
XX  
CC The human apolipoprotein may be produced by a suitable transformed  
CC host, it is effective in treating hyperlipemia and arteriosclerosis.  
XX  
SQ Sequence 267 AA;  
  
Query Match 100.0%; Score 556; DB 7; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-48;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QKLHLEQEKISPLGEEMDRARAHVDALRTHLAPYSDELRORLARLEALKENGARLAE 60  
DB 156 qklhelqekispigeemdrarahvdalrthlapysdelrgrlaarleaalkengarlae 215  
  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLPVLSEFKVSFLSALAEYTKKLTQ 112  
DB 216 yhakatehlstlsekakpaledlrqglpvlsefkvsflsaeeytkkltq 267  
  
RESULT 4  
AAP82128  
ID AAP82128 standard; protein; 267 AA.  
XX  
AC AAP82128;  
XX  
DT 24-OCT-1990 (first entry)  
XX  
DE Entire human preproapoprotein A1.  
XX  
KM human preproapoprotein A1; high density lipoprotein deficiency; ss.  
XX  
OS synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Peptide /label=precursor  
FT Peptide 19..24  
FT Peptide /label=propeptide  
FT Protein 25..267  
FT /label=mature apoprotein  
XX  
PN EP293357-A.  
XX  
PD 30-NOV-1988.  
XX  
PF 24-MAY-1988; 88EP-0870095.  
XX  
PR 28-MAY-1987; 87GB-0012540.  
XX  
PA (UNIO ) UCB SA.  
XX  
PI Bollen A, Gobert J, Wulfert E;  
XX  
DR WPI; 1988-339891/48.  
DR N-PSDB; AAN82064.  
XX  
XX New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures  
XX  
PS Disclosure; ; P; French.  
XX  
CC The cDNA 878bp fragment encoding preproapoprotein A1 was detected  
CC in clone PUB1609 derived from human liver cells.  
CC See also AAN81258.

XX  
SQ Sequence 267 AA;  
  
Query Match 100.0%; Score 556; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-48;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QKLHLEQEKISPLGEEMDRARAHVDALRTHLAPYSDELRORLARLEALKENGARLAE 60  
DB 156 qklhelqekispigeemdrarahvdalrthlapysdelrgrlaarleaalkengarlae 215  
  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLPVLSEFKVSFLSALAEYTKKLTQ 112  
DB 216 yhakatehlstlsekakpaledlrqglpvlsefkvsflsaeeytkkltq 267  
  
RESULT 5  
AAR34032  
ID AAR34032 standard; Protein; 267 AA.  
XX  
AC AAR34032;  
XX  
DT 13-AUG-1993 (first entry)  
XX  
DE Sequence of apo A1.  
XX  
KM Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.  
XX  
OS Homo sapiens.  
XX  
PN WO9307165-A.  
XX  
PD 15-APR-1993.  
XX  
PF 09-OCT-1992; 92WO-US08634.  
XX  
PR 09-OCT-1991; 91US-0774633.  
PR 08-OCT-1992; 92US-055555.  
PR 28-JUN-1992; 92US-0901706.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Curtiss IK, Koduri KR, Smith RS, Witztum JL, Young SG;  
XX  
DR WPI; 1993-134378/16.  
DR N-PSDB; AAQ40030.  
XX  
PT Polypeptide mimic of native apo B-100 and native apo A-I - useful  
PT in assays for LDL and HDL in plasma samples  
XX  
PS Claim 19; pages 105-106; 137pp; English.  
XX  
CC The inventors claim a portion of the polypeptide contg. apo B-100  
CC that immunoreacts with antibodies secreted by the hybridoma MB47  
CC having ATCC Accession No. 8746. Polypeptides specifically claimed  
CC include residues 217-297, 216-310, 216-331, 216-352, 216-377, 1-377,  
CC 205-297, 173-297, 140-297. DNA sequences encoding the polypeptides  
CC are also claimed. Also claimed are a fusion polypeptide that  
CC contains: (a) a first amino acid residue sequence up to 250 residues  
CC in length that includes residues 120-135 of apo A-I, (b) a second  
CC amino acid residue sequence up to 375 residues in length that  
CC includes residues 217-297 of apo B-100 and DNA encoding it.  
XX  
SQ Sequence 267 AA;  
  
Query Match 100.0%; Score 556; DB 14; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-48;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QKLHLEQEKISPLGEEMDRARAHVDALRTHLAPYSDELRORLARLEALKENGARLAE 60  
DB 156 qklhelqekispigeemdrarahvdalrthlapysdelrgrlaarleaalkengarlae 215

Db 156 qklhelgeklspigeemrdrarahvdalrthlapysdelrqlaarlealkengarlae 215

QY 61 YHAKATEHLSLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTNQ 112  
112  
Db 216 yhakatehlsstlsekakpaledlrqglipvlesfkvsflsaeleeytkkltntq 267

RESULT 6  
ID AAR72705 standard; Protein; 267 AA.  
XX AAR72705;  
AC AAR72705;  
XX 31-OCT-1995 (first entry)  
DT 31-OCT-1995 (first entry)  
XX Human apo A-I including signal and propeptide sequences.  
DE Human apo A-I including signal and propeptide sequences.  
XX Apo A-I; LDL cholesterol; low density lipoprotein; lipid.  
KW Apo A-I; LDL cholesterol; low density lipoprotein; lipid.  
XX Homo sapiens.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..18  
FT /label= presignal  
FT 19..24  
FT /label= propeptide  
FT 120..135  
FT /label= claimed  
FT /note= "as part of fusion polypeptide"  
FT 19..240  
FT /label= claimed  
FT /note= "as part of fusion polypeptide"  
XX US5408038-A.  
PN 18-APR-1995.  
PD 18-APR-1995.  
XX 09-OCT-1991; 91US-0774633.  
PF 09-OCT-1991; 91US-0774633.  
XX 09-OCT-1991; 91US-0774633.  
PR 09-OCT-1991; 91US-0774633.  
PR 18-JUN-1992; 92US-0901706.  
PR 08-OCT-1992; 92US-0959946.  
XX (SCRI ) SCRIPPS RES INST.  
XX PA  
XX PI Curtiss LK, Koduri KR, Smith RS, Witztum JL, Young SG;  
PI N-PSDB; AAQ89634.  
DR WPI; 1995-161146/21.  
DR N-PSDB; AAQ89634.  
XX New apo:lipoprotein B-100 peptide(s) and fusion peptide(s) - used  
PT in assay systems for detecting LDL and HDL cholesterol levels in  
PT body fluids.  
XX Claim 10; Fig 2; 41pp; English.  
XX AAA89634 and AAR72705 depict the AA sequence of human apo A-I and its  
CC corresp. cDNA, including presignal residues and propeptide  
CC residues, according to Seilhamer et al., DNA 3(4):309 (1984).  
CC A dispersible apo A-I/B-100 fusion polypeptide is claimed which  
CC contains a first AA sequence of apo A-I and that includes at  
CC least AA sequence positions 120-135 (see AAR72606) and which reacts  
CC with pan anti-apo A-I antibodies such as: AI-4 ATCC HB8744; AI-7  
CC ATCC HB 8745; AI-9 ATCC HB 8741; AI-10 ATCC HB 9200; AI-11 ATCC  
CC HB 9201; AI-12 ATCC HB 9202; AI-13 ATCC HB 9203; AI-14 ATCC HB  
CC 9204; AI-18 ATCC HB 9507.  
XX SQ Sequence 267 AA;

Query Match 100.0%; Score 556; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-48;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKLHELOEKSLPGEEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
112  
Db 156 qklhelgeklspigeemrdrarahvdalrthlapysdelrqlaarlealkengarlae 215

QY 61 YHAKATEHLSLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTNQ 112  
112  
Db 216 yhakatehlsstlsekakpaledlrqglipvlesfkvsflsaeleeytkkltntq 267

RESULT 7  
ID AAY18675 standard; Protein; 267 AA.  
XX AAY18675;  
AC AAY18675;  
XX 09-JUL-1999 (first entry)  
DT 09-JUL-1999 (first entry)  
XX Human apolipoprotein AI protein sequence.  
DE Human apolipoprotein AI protein sequence.  
XX Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.  
XX Homo sapiens.  
OS Homo sapiens.  
XX WO916409-A2.  
PN WO916409-A2.  
XX 08-APR-1999.  
PD 08-APR-1999.  
XX 28-SEP-1998; 98WO-US20329.  
PF 28-SEP-1998; 98WO-US20329.  
XX 29-SEP-1997; 97US-0940136.  
PR 29-SEP-1997; 97US-0940136.  
XX (BUTT/) BUTTNER K.  
PA (CORN/) CORNUT I.  
PA (DASS/) DASSEUX J.  
PA (DUFO/) DUFOURCQ J.  
PA (METZ/) METZ G.  
PA (SEK/) SEKUL R.  
XX Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;  
PI Sekul R;  
PI N-PSDB; AAX55971.  
DR WPI; 1999-254921/21.  
DR N-PSDB; AAX55971.  
XX Nucleic acid encoding apolipoprotein A-I agonist peptides  
XX Example; Fig 1; 232pp; English.  
XX The present invention describes a nucleic acid (A) encoding an  
CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,  
CC which forms an amphipathic alpha-helix in presence of lipids. (A),  
CC optionally as a complex with lipids, and host cells that contain (A),  
CC are useful for gene therapy, or prevention, of diseases associated with  
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I  
CC deficiency, hypertriglyceridemia and metabolic syndrome, also to treat  
CC endotoxemia (septic shock). Host cells containing (A) can also be used  
CC to study the role of apoA-I in lipid metabolism. (B) can be used  
CC diagnostically, e.g. to measure serum HDL (particularly its  
CC subpopulation involved in retrograde cholesterol transport) and for  
CC imaging the circulatory system or HDL accumulations at fatty streaks.  
XX The present sequence represents human apoA-I.  
XX SQ Sequence 267 AA;

Query Match 100.0%; Score 556; DB 20; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-48;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 OKLHELQEKLSPLGEEMDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGARLAE 60
        |||||||
Db      156 qklhelqeklsplgeemdrarahvdalrthlapydsdelqrlaarlalalkengarlae 215
QY      61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKLNTQ 112
        |||||||
Db      216 yhakatehlstlsekakpaledlrqglpylesfkvsflsalseeytklntq 267

RESULT  8
AAB47620
ID      AAB47620 standard; Protein; 267 AA.
XX
AC      AAB47620;
XX
DT      21-JAN-2002 (first entry)
XX
DE      Full length Apo-A1.
XX
KW      Apolipoprotein; Apo-A1; Apo-A-I fragment T-cell activation inhibitor;
KW      AFTI; monocyte; IL-1; interleukin 1; TNF; tumour necrosis factor;
KW      acute pancreatitis; Alzheimer's disease; asthma; cancer; fever;
KW      inflammatory bowel disease; ischemia; multiple sclerosis; osteoporosis;
KW      Parkinson's disease; psoriasis; probe.
XX
OS      Homo sapiens.
XX
FH      Key
FT      Binding-site      Location/Qualifiers
                        44..65
FT      Binding-site      /label= Helical lipid binding domain
                        220..241
FT      Binding-site      /label= Helical lipid binding domain
                        74..111
FT      Binding-site      /note= "Involved in lipoprotein-mediated cholesterol
                        efflux from monocytes"
FT      Binding-site      149..219
FT      Binding-site      /label= Receptor binding domain
FT      Binding-site      99..120
FT      Binding-site      /label= Major antigenic epitope domain
FT      Binding-site      99..143
FT      Binding-site      /label= Hinged domain
FT      Binding-site      66..120
FT      Binding-site      /label= Phylogenetically conserved domain
FT      Binding-site      90..111
FT      Binding-site      /note= "Involved in lectin-cholesterol acyltransferase
                        activity"
FT      Binding-site      44..65
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      66..98
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      99..120
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      121..142
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      143..164
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      165..208
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      209..219
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      220..241
FT      Binding-site      /label= Amphipathic helix
FT      Binding-site      25..194
FT      Binding-site      /label= AFTI
FT      Binding-site      /note= "18 kD N-terminal fragment"
FT      Binding-site      25..144
FT      Binding-site      /label= AFTI
FT      Binding-site      /note= "13 kD N-terminal fragment"
FT      Binding-site      156..267
FT      Binding-site      /label= AFTI
FT      Binding-site      /note= "13 kD C-terminal fragment"
XX
PN      WO200168852-A2.
```

```
XX      XX
PD      20-SEP-2001.
XX      XX
PF      13-MAR-2001; 2001WO-US07826.
XX      PR
XX      13-MAR-2000; 2000US-189008P.
XX      PA
XX      (AMGE-) AMGEN INC.
XX      PI
PI      Edwards CK, Burger D, Dayer J, Kohno T;
XX      DR
DR      WPI; 2001-596908/67.
DR      DR
DR      N-PSDB; AAH43623.
XX      PT
PT      Apo-A-I fragment T-cell activation inhibitor (AFTI) polynucleotides,
PT      useful for treating, diagnosing, ameliorating diseases associated with
PT      IL-1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease
PT      and asthma -
XX      PS
PS      Claim 1; Fig 1A; 132pp; English.
XX      CC
CC      This sequence shows full length apolipoprotein (Apo-A1). Fragments
CC      of Apo-A1 may be used as Apo-A-I fragment T-cell activation inhibitors
CC      (AFTI). These fragments are selected from an 18 kD N-terminal fragment
CC      (amino acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144)
CC      and a 13 kD C-terminal fragment (amino acids 156-267). The AFTI
CC      polypeptides and polynucleotides are useful for regulating T-cell
CC      mediated activation of monocytes and for treating, diagnosing,
CC      ameliorating diseases associated with IL-1 and/or TNF activity.
CC      The diseases are acute pancreatitis, Alzheimer's disease, asthma,
CC      cancer, fever, inflammatory bowel disease, ischemia, multiple
CC      sclerosis, osteoporosis, Parkinson's disease or psoriasis. Numerous
CC      examples of other diseases are given in the specification.
CC      The AFTI nucleic acids are useful as hybridization probes in diagnostic
CC      assays to test for the presence of an AFTI or Apo-A-I DNA in mammalian
CC      tissue or bodily fluid samples.
XX      SQ
SQ      Sequence      267 AA;

Query Match      100.0%; Score 556; DB 22; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.1e-48;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 OKLHELQEKLSPLGEEMDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGARLAE 60
        |||||||
Db      156 qklhelqeklsplgeemdrarahvdalrthlapydsdelqrlaarlalalkengarlae 215
QY      61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKLNTQ 112
        |||||||
Db      216 yhakatehlstlsekakpaledlrqglpylesfkvsflsalseeytklntq 267

RESULT  9
AAP80668
ID      AAP80668 standard; protein; 268 AA.
XX
AC      AAP80668;
XX
DT      24-OCT-1990 (first entry)
XX
DE      Recombinant human preproapoprotein A1.
XX
KW      human preproapoprotein A1; high density lipoprotein deficiency; ss.
XX      OS
OS      synthetic.
XX
FH      Key
FT      Peptide      Location/Qualifiers
                        1..18
FT      Peptide      /label=precursor
FT      Peptide      19..25
FT      Peptide      /label=proprotein
FT      Protein      26..268
```

FT /label=mature apoprotein A1  
XX EP293357-A.  
PN 30-NOV-1988.  
XX 24-MAY-1988; 88EP-0870095.  
XX 28-MAY-1987; 87GB-0012540.  
XX (UNIO ) UCB SA.  
XX Bollen A, Gobert J, Wulfert E;  
XX WPI; 1988-339891/48.  
XX N-PSDB; AAN81258.  
XX  
XX New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures  
XX  
XX Claim 1; Page 12; 25pp; French.  
PS Met at posn 19 is inserted as an extra amino acid c.f. wild-type  
XX  
XX protein.  
CC The DNA fragment used to replace the wild-type sequence encoding  
CC amino acids -6 to +14 (= Arg 20 to Leu 39) modifies codons  
CC corresponding to wild-type amino acids -6, -1, 1, 3, 4, 5, 6, 7,  
CC 10, 11 and 14. The changed codons still encode the same amino acids  
CC as in the wild-type protein but reduce formation of secondary  
CC structures in mRNA.  
CC See also AAN82064.  
XX  
SQ Sequence 268 AA;  
  
Query Match 100.0%; Score 556; DB 9; Length 268;  
Best Local Similarity 100.0%; Pred. No. 1.1e-48;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QKHLHLEKLSPLGEMRDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
Db 157 qklhelqeklsplgeemrdrarahvdalrthlapysdelrqlaarlealkengarlae 216  
  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESEKVSFLSALLEEYTKKLTNQ 112  
Db 217 yhakatehlstlsekakpaledlrqglpvlsefkvsflsaleeytkkltntq 268  
  
RESULT 10  
AAW08602  
ID AAW08602 standard; Protein; 267 AA.  
XX  
AC AAW08602;  
XX  
DT 04-SEP-1997 (first entry)  
XX  
DE Human apolipoprotein A-1 variant "Paris" protein sequence.  
XX  
XX Human; apolipoprotein A-1; variant; mutation; serum lipid; cholesterol;  
KW high density lipoprotein; triglyceride; symptom; cardiovascular disease;  
KW atherosclerosis; restenosis; myocardial infarct; angina; death; stroke;  
KW cardiac decompensation; metabolic deficit.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..24  
FT /note= "prepro peptide"  
FT Region 172..178  
FT /note= "this sequence which contains the mutated amino  
FT acid residue is claimed - claim 2"  
FT Misc-difference 175  
FT /note= "changed from Arg residue in wild type protein,"

FT due to a C to T transition mutation"  
XX  
XX WO9637608-A1.  
PN 28-NOV-1996.  
XX  
XX 20-MAY-1996; 96WO-FR00747.  
XX 22-MAY-1995; 95FR-0006061.  
XX  
XX (INSP ) INST PASTEUR LILLE.  
PA (RHON ) RHONE POULENC RORER SA.  
PA (UYPA-) UNIV CURIE PARIS VI P & M.  
XX  
XX Benoit P, Bruckert E, Denefle P, Duverger N, Fruchart J;  
PI Luc G, Turping, Assmann G, Funke H;  
PI  
XX WPI; 1997-021218/02.  
XX N-PSDB; AAT43691.  
XX  
XX New variant of human apoA-1 with Cys at position 151 - has  
PT anti-atherogenic activity for treatment and prevention of  
PT cardiovascular disease  
XX  
XX Claim 2; Page -; 58pp; French.  
PS  
XX  
XX This is the amino acid sequence of a human apolipoprotein A-1 variant  
XX designated the "Paris" variant which has a Cys replacing the Arg residue  
XX at position 151. The substitution is generated by a mutation of  
XX the C nucleotide at position 523 in the wild type gene to a T residue,  
XX changing the encoded residue from an Arg to a Cys. The gene was isolated  
XX from a patient with an unusual pattern of serum lipids i.e. low levels of  
XX apoA-1 and high density lipoprotein (HDL)-cholesterol and high  
XX triglycerides, but showing symptoms of atherosclerosis. The new variant  
XX protein has anti-atherogenic activity so is useful for treatment and  
XX prevention of cardiovascular diseases such as atherosclerosis,  
XX restenosis, myocardial infarct, angina, sudden death, stroke and cardiac  
XX decompensation, or more generally any condition involving genetic or  
XX metabolic deficit of apoA-1.  
XX Note: this sequence is not given in the specification but is generated  
XX from the wild type apoA-1 gene disclosed in the specification and has  
XX the appropriate amino acid changed.  
XX  
SQ Sequence 267 AA;  
  
Query Match 98.6%; Score 548; DB 18; Length 267;  
Best Local Similarity 99.1%; Pred. No. 7.1e-48;  
Matches 111; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 QKHLHLEKLSPLGEMRDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
Db 156 qklhelqeklsplgeemrdrarahvdalrthlapysdelrqlaarlealkengarlae 215  
  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESEKVSFLSALLEEYTKKLTNQ 112  
Db 216 yhakatehlstlsekakpaledlrqglpvlsefkvsflsaleeytkkltntq 267  
  
RESULT 11  
AAR56863  
ID AAR56863 standard; Protein; 264 AA.  
XX  
AC AAR56863;  
XX  
DT 26-JAN-1995 (first entry)  
XX  
DE Apo-lipoprotein AI-M.  
XX  
XX Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid pKP683;  
KW vector; atherosclerosis; cardiovascular disease.  
XX  
OS Homo sapiens.



XX PN W09413819-A.  
XX PD 23-JUN-1994.  
XX PF 09-DEC-1993; 93WO-SE01061.  
XX PR 11-DEC-1992; 92SE-0003753.  
XX PA (KABI ) KABI PHARMACIA AB.  
XX PI Abrahmsen L, Holmgren E, Kaldereen C, Lake M, Mikaelsson A;  
XX PI Sejlitz T;  
XX DR WPI; 1994-217892/26.  
XX DR N-PSDB; AAQ68357.

PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 3; 33pp; English.

CC Plasmid PKP683 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of PKP683 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.

XX SQ Sequence 264 AA;

Query Match 98.0%; Score 545; DB 15; Length 264;  
Best Local Similarity 98.2%; Pred. No. 1.4e-47;  
Matches 110; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKLHLEQKLSPLGEEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
DB 153 qklhelqeklsplgeemdrarahvdalrthlapysdelrqlaarlealkengarlar 212

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLPVLESEFKVSFLSALAEYTKKLTQ 112  
DB 213 yhakatehlstlsekakpaledlrqglpvlesfrvsflsaeeytkkltq 264

#### RESULT 12

AAR56864  
ID AAR56864 standard; Protein; 264 AA.

XX AC AAR56864;

DT 26-JAN-1995 (first entry)

DE Apo-lipoprotein AI-M.

KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid PKP764;  
KW vector; atherosclerosis; cardiovascular disease.

OS Homo sapiens.

PN W09413819-A.

PD 23-JUN-1994.

PF 09-DEC-1993; 93WO-SE01061.

PR 11-DEC-1992; 92SE-0003753.

PA (KABI ) KABI PHARMACIA AB.

PI Abrahmsen L, Holmgren E, Kaldereen C, Lake M, Mikaelsson A;  
PI Sejlitz T;

DR WPI; 1994-217892/26.  
DR N-PSDB; AAQ68358.

PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 4; 33pp; English.

CC Plasmid PKP764 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of PKP764 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.

XX SQ Sequence 264 AA;

Query Match 98.0%; Score 545; DB 15; Length 264;  
Best Local Similarity 98.2%; Pred. No. 1.4e-47;  
Matches 110; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKLHLEQKLSPLGEEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
DB 153 qklhelqeklsplgeemdrarahvdalrthlapysdelrqlaarlealkengarlar 212

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLPVLESEFKVSFLSALAEYTKKLTQ 112  
DB 213 yhakatehlstlsekakpaledlrqglpvlesfrvsflsaeeytkkltq 264

#### RESULT 13

AAU33170  
ID AAU33170 standard; Protein; 299 AA.

XX AC AAU33170;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3661.

KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN W0200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -

PS Claim 20; Page 718; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

XX Sequence 299 AA;

Query Match 93.9%; Score 522; DB 22; Length 299;

Best Local Similarity 92.9%; Pred. No. 3.6e-45;

Matches 104; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 OKLHELOEKLSPLGEMRDRARAHVDALRTHLAPYSDELRORLARLEALKENGARLAE 60

Db 188 qklhelgelsplgeemrdrarahvdalrthlapysdelrqlgarlgallrengarmg 247

OY 61 YHAKATEHLSTLSEKAKPALEDLROGLLPVLESFKVSFLSALAEYTKKLTNQ 112

Db 248 yhaqatehstlsekakpaledlrgllpvlesfkvsflsalseytkkltntq 299

RESULT 14

AAU30470

ID AAU30470 standard; Protein; 359 AA.

AC AAU30470;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #961.

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy -

XX Claim 20; Page 297; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically

CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

XX Sequence 359 AA;

Query Match 89.3%; Score 496.5; DB 22; Length 359;

Best Local Similarity 88.5%; Pred. No. 1.8e-42;

Matches 100; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

OY 1 OKLHEL-OKISPLGEMRDRARAHVDALRTHLAPYSDELRORLARLEALKENGARLA 59

Db 247 eslhelarlsplgeemrdrarahvdalrthlapysdelrqlgarlgallrengarmg 306

OY 60 EYHAKATEHLSTLSEKAKPALEDLROGLLPVLESFKVSFLSALAEYTKKLTNQ 112

Db 307 qyhaqatchstlsekakpaledlrgllpvlesfkvsflsalseytkkltntq 359

RESULT 15

AAU30268

ID AAU30268 standard; Protein; 318 AA.

AC AAU30268;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #759.

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy -

XX Claim 20; Page 270; 765pp; English.

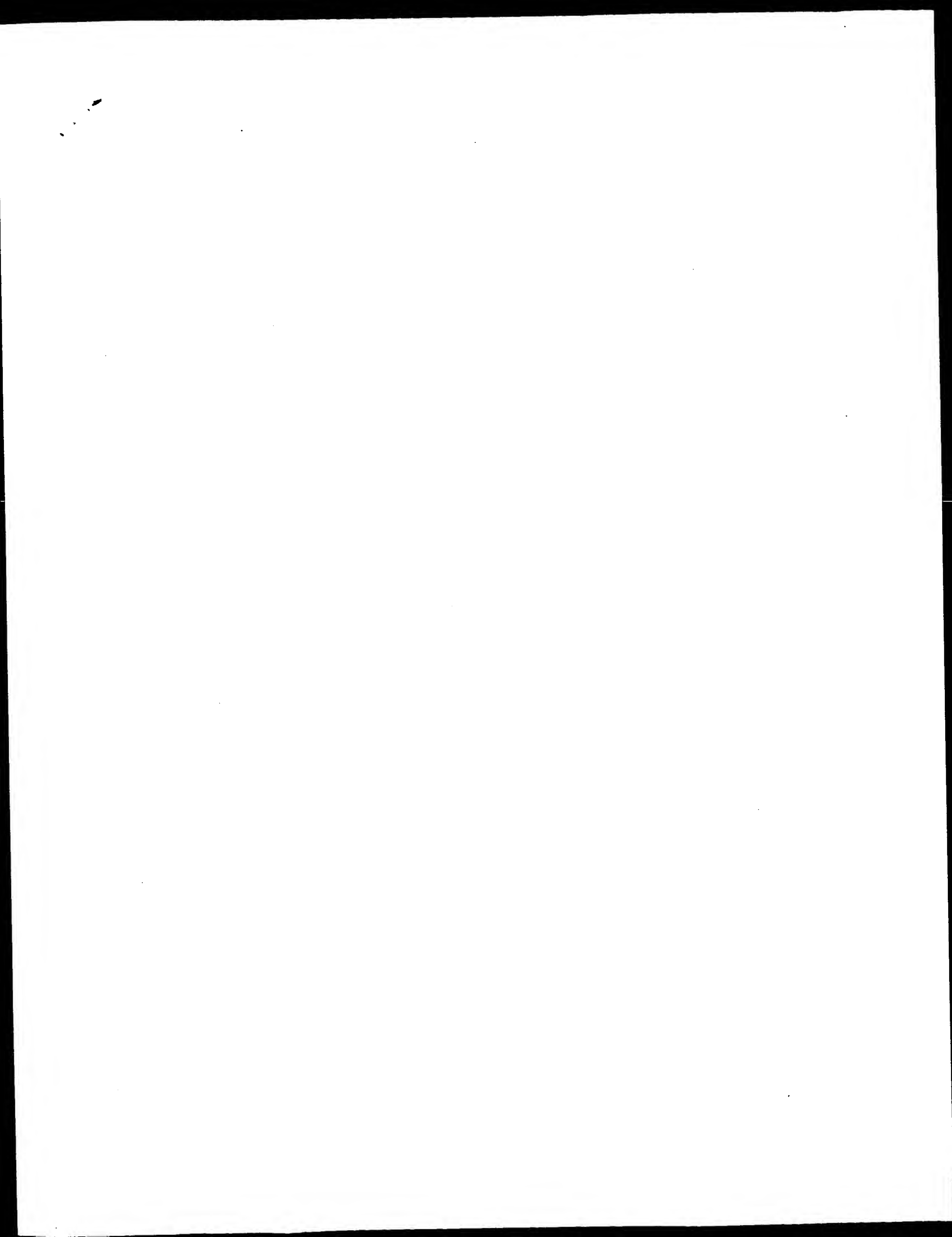
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to

CC 'increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AA029510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 318 AA;

Query Match	73.7%;	Score 409.5;	DB 22;	Length 318;
Best Local Similarity	77.4%;	Pred. No. 1.1e-33;		
Matches	89;	Conservative	8;	Mismatches 15; Indels 3; Gaps 3;

```
QY 1 QKIHLEL-QEKISPLGEEEM-RDRARAHVDALRTHLAPYSDELQRQLA-ARLEALKENGAR 57
   : :|:|:| :|:|:|:| :| | | | | | | | | | | :| | | | | | | |
Db 179 eslhelarstlsplgeavsrprarpmwdalrthlapysdemmalgraplgalrengar 238
QY 58 LAEYHAKATEHLUSTLSEKAKPALEDLRQGLPVLESFKVSLALEEYTKKLTNQ 112
   : :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 239 mgqyhaqatehltstlsekakpaledlrqglpvlesfkvsfalsaleeytkkltncq 293
```

Search completed: September 22, 2002, 12:05:16  
Job time: 291 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:02:51 ; Search time 51.49 Seconds  
(without alignments)  
53.130 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_156\_267

Perfect score: 556  
Sequence: 1 QKHELEKLSPLGEMRDR.....SFKVSFLALEYTKKLTNQ 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	556	100.0	267	1	US-07-959-946-3
2	556	100.0	267	1	US-08-333-577-3
3	556	100.0	267	1	US-08-952-796-2
4	556	100.0	267	5	PCT-US92-08634-3
5	548	98.6	200	4	US-08-952-796-15
6	548	98.6	264	1	US-08-448-606-6
7	92	16.5	317	1	US-07-709-949-2
8	90	16.2	317	4	US-08-949-155-6
9	86	15.5	64	2	US-08-292-870-1
10	83	14.9	1184	4	US-09-541-782-2
11	80.5	14.5	110	1	US-07-849-389-7
12	80.5	14.5	220	2	US-08-726-306A-29
13	75.5	13.6	878	3	US-08-941-936-2
14	75.5	13.6	1939	4	US-09-310-187A-1
15	73.5	13.2	1886	4	US-08-938-105-3
16	73	13.1	1618	1	US-07-853-913-4
17	72.5	13.0	618	3	US-08-834-306-65
18	72.5	13.0	618	4	US-08-993-674A-65
19	72	12.9	1561	3	US-08-894-017-23
20	72	12.9	1565	6	5352450-2
21	71.5	12.9	955	1	US-08-006-676B-1
22	71.5	12.9	955	1	US-08-282-845-2
23	71.5	12.9	955	2	US-08-428-414A-3
24	71.5	12.9	955	5	PCT-US94-00324-1
25	69.5	12.5	2101	1	US-08-466-390-4
26	69.5	12.5	2101	1	US-08-470-950-4
27	69.5	12.5	2101	1	US-08-467-781-4

28	69.5	12.5	2101	1	US-08-195-487-4	Sequence 4, Appli
29	69.5	12.5	2101	2	US-08-483-924-4	Sequence 4, Appli
30	69.5	12.5	2101	4	US-09-452-294-1	Sequence 1, Appli
31	69.5	12.5	2101	5	PCT-US93-06160-4	Sequence 4, Appli
32	69	12.4	566	2	US-08-533-669A-8	Sequence 8, Appli
33	69	12.4	566	2	US-08-511-872-2	Sequence 2, Appli
34	68.5	12.3	337	1	US-08-445-135-4	Sequence 4, Appli
35	68.5	12.3	339	1	US-08-433-854-4	Sequence 4, Appli
36	68.5	12.3	339	1	US-08-174-745A-4	Sequence 4, Appli
37	68.5	12.3	339	2	US-08-195-947-4	Sequence 4, Appli
38	68.5	12.3	339	2	US-08-433-885-4	Sequence 4, Appli
39	68.5	12.3	339	2	US-08-433-908B-4	Sequence 4, Appli
40	68.5	12.3	339	4	US-08-410-614-4	Sequence 4, Appli
41	68	12.2	180	1	US-08-328-254-7	Sequence 7, Appli
42	68	12.2	184	2	US-08-715-204-5	Sequence 5, Appli
43	68	12.2	184	2	US-08-691-814B-50	Sequence 50, Appli
44	68	12.2	184	3	US-09-162-597-5	Sequence 5, Appli
45	68	12.2	337	1	US-08-317-223-3	Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-07-959-946-3  
; Sequence 3, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Wiltzium, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959, 946  
; FILING DATE: 19921008  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901, 706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamsen, Edward P.  
; REGISTRATION NUMBER: 29,381  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-946-3  
Query Match 100.0%; Score 556; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 2.4e-53;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
|||||  
Db 156 QKLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 215

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTQ 112  
|||||  
Db 216 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTQ 267

## RESULT 2

US-08-333-577-3

; Sequence 3, Application US/08333577  
; Patent No. 5786206

## ; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.

; APPLICANT: Koduri, Raju

; APPLICANT: Young, Stephen G.

; APPLICANT: Wiltum, Joseph L.

; APPLICANT: Curtiss, Linda K.

; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a

; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens

; NUMBER OF SEQUENCES: 20

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &amp;

; ADDRESSEE: Milnamow, Ltd.

; STREET: 180 No. 5786206th Stetson, Suite 4700

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60601

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/333,577

; FILING DATE:

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Gamsen, Edward P.

; REGISTRATION NUMBER: 29,381

; REFERENCE/DOCKET NUMBER: SCRF 234.0

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312)616-5400

; TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-333-577-3

Query Match 100.0%; Score 556; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.4e-53;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
|||||  
Db 156 QKLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 215

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTQ 112  
|||||  
Db 216 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTQ 267

## RESULT 3

US-08-952-796-2

; Sequence 2, Application US/08952796  
; Patent No. 6258596

## ; GENERAL INFORMATION:

; APPLICANT: BENNETT, Patrick

; APPLICANT: BRUCKERT, Eric

; APPLICANT: DENEFFLE, Patrice

; APPLICANT: DUBERGER, Nicolas

; APPLICANT: FRUCHART, Jean-Charles

; APPLICANT: LUC, Gerald

; APPLICANT: TURPIN, Gerard

; APPLICANT: ASSMANN, Gerd

; APPLICANT: FUNKE, Harald

; TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPROTEIN A-I

; NUMBER OF SEQUENCES: 19

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Rhone-Poulenc Rorer Inc.

; STREET: 500 Arcola Road, Mailstop 3C43

; CITY: Collegeville

; STATE: PA

; COUNTRY: USA

; ZIP: 19426

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/952,796

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 95/06061

; FILING DATE: 22-MAY-1995

; APPLICATION NUMBER: WO FR96/00747

; FILING DATE: 20-MAY-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fehner Esq., Paul F.

; REGISTRATION NUMBER: 35,135

; REFERENCE/DOCKET NUMBER: ST95031-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (610) 454-3839

; TELEFAX: (610) 454-3808

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-952-796-2

Query Match 100.0%; Score 556; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.4e-53;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 60  
|||||  
Db 156 QKLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAE 215

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTQ 112  
|||||  
Db 216 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALAEYTKKLTQ 267

## RESULT 4

PCT-US92-08634-3

; Sequence 3, Application PC/TUS9208634  
; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.

; APPLICANT: Koduri, Raju

; APPLICANT: Young, Stephen G.

; APPLICANT: Wiltum, Joseph L.

APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 North Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08634  
FILING DATE: 19921009  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gansson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-08634-3

Query Match 100.0%; Score 556; DB 5; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.4e-53;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELRLQRLARLEALKENGARLAE 60  
|||||  
Db 156 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELRLQRLARLEALKENGARLAE 215  
|||||

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112  
|||||  
Db 216 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 267  
|||||

RESULT 5  
US-08-952-796-15  
; Sequence 15, Application US/08952796  
; Patent No. 6258596  
; GENERAL INFORMATION:  
; APPLICANT: BENOIT, Patrick  
; APPLICANT: BRUCKERT, Eric  
; APPLICANT: DENEFFE, Patrice  
; APPLICANT: DUBERGER, Nicolas  
; APPLICANT: FRUCHART, Jean-Charles  
; APPLICANT: LUC, Gerald  
; APPLICANT: TURPIN, Gerard  
; APPLICANT: ASSMANN, Gerd  
; APPLICANT: FUNKE, Harald  
; TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, Mailstop 3C43  
; CITY: Collegeville

STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehner Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 200 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-15

Query Match 98.6%; Score 548; DB 4; Length 200;  
Best Local Similarity 99.1%; Pred. No. 1.2e-52;  
Matches 111; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELRLQRLARLEALKENGARLAE 60  
|||||  
Db 89 QKLHLEQEKLSPLGEMDRARAHVDALRTHLAPYSDELRLQRLARLEALKENGARLAE 148  
|||||

QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 112  
|||||  
Db 149 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLTQ 200  
|||||

RESULT 6  
US-08-448-606-6  
; Sequence 6, Application US/08448606  
; Patent No. 5721114  
; GENERAL INFORMATION:  
; APPLICANT: Abrahams n, Lars  
; APPLICANT: Holmgren, Erik  
; APPLICANT: Kalder n, Christina  
; APPLICANT: Lake, Mats  
; APPLICANT: Mikaelsson, sa  
; APPLICANT: Sejlitz, Torsten  
; TITLE OF INVENTION: Expression System For Producing  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pollock, Vande Sande & Priddy  
; STREET: 1990 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.  
; ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,606  
; FILING DATE: 25-AUG-1995  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/SE93/01061  
; FILING DATE: 09-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9203753-0  
; FILING DATE: 11-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amernick, Burton A.  
; REGISTRATION NUMBER: 24,852  
; REFERENCE/DOCKET NUMBER: 0151/00121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)331-7111  
; TELEFAX: (202) 293-6229  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-448-606-6

Query Match 98.6%; Score 548; DB 1; Length 264;  
Best Local Similarity 99.1%; Pred. No. 1.8e-52;  
Matches 111; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKLHELQKLSPLGEMDRARAHVDALETHLAPYSDELRLRLAALKEKNGARLAE 60  
Db 153 QKLHELQKLSPLGEMDRARAHVDALETHLAPYSDELRLRLAALKEKNGARLAE 212  
QY 61 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEETKKLNTQ 112  
Db 213 YHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEETKKLNTQ 264

RESULT 7  
US-07-709-949-2  
; Sequence 2, Application US/07709949  
; Patent No. 5472858  
; GENERAL INFORMATION:  
; APPLICANT: Attie, Alan D  
; APPLICANT: Gretch, Daniel G  
; APPLICANT: Sturley, Stephen L  
; APPLICANT: Beckage, Nancy E  
; TITLE OF INVENTION: Production of Recombinant Proteins in  
; TITLE OF INVENTION: Insect Larvae  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: P.O. Box 2113  
; CITY: Madison  
; STATE: WI  
; COUNTRY: USA  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/709,949  
; FILING DATE: 19910604  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seay, Nicholas J  
; REGISTRATION NUMBER: 27,386  
; REFERENCE/DOCKET NUMBER: 9629691801  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (608) 251-5000

TELEFAX: (608)251-9166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 317 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-709-949-2

Query Match 16.5%; Score 92; DB 1; Length 317;  
Best Local Similarity 28.9%; Pred. No. 0.02;  
Matches 35; Conservative 20; Mismatches 30; Indels 36; Gaps 4;

QY 5 ELQKLSPLGEMDR-----ARAHVDA-----LRLHAPYSDELRLR 42  
Db 95 ELEEQLPVAETTRARLSKELQTAQARLGADMEDVCGRLVQYRGVEQAMLGQSTELRVR 154  
QY 43 LARLEALK-----ENGARLAETHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFK 95  
Db 155 LASHLRKLRKRLRPDDLQKRLAVYQAGA-----REGAERGLSAIRERLGPLVEQGR 207  
QY 96 V 96  
Db 208 V 208

RESULT 8  
US-08-949-155-6  
; Sequence 6, Application US/08949155  
; Patent No. 6271436  
; GENERAL INFORMATION:  
; APPLICANT: Piedrahita, Jorge A  
; APPLICANT: Bazer, Fuller W  
; TITLE OF INVENTION: Compositions and Methods for the  
; TITLE OF INVENTION: Generation of Transgenic Animal Species  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARNOLD, WHITE AND DURKEE  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TX  
; COUNTRY: US  
; ZIP: 77210-4433  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/949,155  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/027,338  
; FILING DATE: 11-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/046,094  
; FILING DATE: 09-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hibler, David W.  
; REGISTRATION NUMBER: 41,071  
; REFERENCE/DOCKET NUMBER: TMAK:177  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 317 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-949-155-6





QY	105	YT	106
		:	1
Db	608	FT	609

```

RESULT 14
US-09-310-187A-1
; Sequence 1, Application US/09310187A
; Patent No. 6358751
; GENERAL INFORMATION:
; APPLICANT: Benichou, Gilles
; APPLICANT: Fedoseyeva, Eugenia
; TITLE OF INVENTION: Involvement of Autoantigens in Cardiac
; TITLE OF INVENTION: Graft Rejection
; FILE REFERENCE: UCSF-090
; CURRENT APPLICATION NUMBER: US/09/310,187A
; CURRENT FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-310-187A-1

```

Query Match	13.6%;	Score 75.5;	DB 4;	Length 1939;
Best Local Similarity	27.0%;	Pred. No. 14;		
Matches 37; Conservative	21;	Mismatches 44;	Indels 35;	Gaps 6;

QY	1	ÖKLHELQÖELSLPLGEEEMRDR--ARAHVDALRTTHIAPYSDELROQLAAREALKENGGA--	56
	:	:    : :     :	:   :
Db	1108	KKIKENQARIIEELELEAEERTARA KVEKLRSDLSRELFEISER-----LEEAGGATS	1160
QY	57	-----RLAEYH-----AKAT-EHLSTLSEKAKPALE-----DLRÖGLLPVLE	92
	:	: :     :   :   :	:   :
Db	1161	VÖIEMNKKREAEFQÖKMRRDLEAEATLQHEATAALUKKHADSVAEELGEÖIDNLÖRVKÖKLE	1220
QY	93	SFKVSFLSALLEEYTKKL	109
	:	: :   :	
Db	1221	KEKSEEKLELDVDTSNM	1237

```

1  RESULT 15
2  US-08-938-105-3
3  ; Sequence 3, Application US/08938105
4  ; Patent No. 6353151
5  ;
6  ; GENERAL INFORMATION:
7  ;
8  ; APPLICANT: Ielnwand, Leslie A.
9  ; APPLICANT: Vikstrom, Karen L.
10 ; TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
11 ;
12 ; NUMBER OF SEQUENCES: 3
13 ;
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: Sheridan Ross P.C.
16 ; STREET: 1700 Lincoln St., Suite 3500
17 ;
18 ; CITY: Denver
19 ;
20 ; STATE: CO
21 ;
22 ; COUNTRY: U.S.A.
23 ;
24 ; ZIP: 80203
25 ;
26 ; COMPUTER READABLE FORM:
27 ;
28 ; MEDIUM TYPE: Floppy disk
29 ;
30 ; COMPUTER: IBM PC compatible
31 ;
32 ; OPERATING SYSTEM: PC-DOS/MS-DOS
33 ;
34 ; SOFTWARE: Patentln Release #1.0, Version #1.30
35 ;
36 ; CURRENT APPLICATION DATA:
37 ; APPLICATION NUMBER: US/08/938,105
38 ;
39 ; FILING DATE:
40 ;
41 ; CLASSIFICATION:
42 ;
43 ; ATTORNEY/AGENT INFORMATION:
44 ;
45 ; NAME: Crook, Wannell M.
46 ;
47 ; REGISTRATION NUMBER: 31,071
48 ;
49 ; REFERENCE/DOCKET NUMBER: 3595-4
50 ;
51 ; TELECOMMUNICATION INFORMATION:
52 ;
53 ; TELEPHONE: (303) 863-9700
54 ;
55 ; TELEFAX: (303) 863-0223
56 ;
57 ; INFORMATION FOR SEQ ID NO: 3:
58 ;
59 ; SEQUENCE CHARACTERISTICS:
60 ;
61 ; LENGTH: 1886 amino acids
62 ;

```

```

;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-938-105-3

```

Query Match	13.28;	Score 73.5;	DB 4;	Length 1886;
Best Local Similarity	27.08;	Pred. No. 22;		
Matches 37; Conservative	20;	Mismatches 45;	Indels 35;	Gaps 6;

QY	1	QKHLHÖEKISPLGCEMRDR--	ARAHVDALRTHLAPYSDELQRÖRLAARLEALKENGGA--	56
		:      :  :  :      :  :  :		
Db	1055	KKLKENQARIIEELLEELEAERTARAKVEKLRSIDLTRLEEISER-----	LEBAGGATS	1107
QY	57	-----RLAEYH-----	AKAT-EHLSTLSEKAPALE-----	DLRÖGLLPVLE 92
		:  :  :  :  :  :  :		
Db	1108	VÖIEMNKKREAEFQÖKMRDLEAATLÖHEATAALRKKHADSVAEALGEOIDNLÖRVKÖKLE		1167
QY	93	SFKVSFLSALLEEYTKKL	109	
		:  :  :		
Db	1168	KEKSEKLELDDVTSHM	1184	

Search completed: September 22, 2002, 12:02:52  
Job time: 252 sec

Sun Sep 22 12:10:09 2002

us-09-803-918a-2\_copy\_156\_267.ra

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:02:50 : Search time 51.49 Seconds  
(without alignments)  
56.925 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_25\_144  
Perfect score: 631  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....LDDFQKKWQEMELYRQKVE 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	631	100.0	264	1	US-08-448-606-6 Sequence 6, Appli
2	631	100.0	267	1	US-07-959-946-3 Sequence 3, Appli
3	631	100.0	267	1	US-08-333-577-3 Sequence 3, Appli
4	631	100.0	267	4	US-08-952-796-2 Sequence 2, Appli
5	631	100.0	267	5	PCT-US92-08634-3 Sequence 3, Appli
6	404	64.0	200	4	US-08-952-796-15 Sequence 15, Appli
7	198	31.4	44	2	US-08-292-870-2 Sequence 2, Appli
8	188	29.8	64	2	US-08-292-870-1 Sequence 1, Appli
9	138	21.9	32	2	US-08-292-870-3 Sequence 3, Appli
10	124	19.7	25	2	US-08-292-870-4 Sequence 4, Appli
11	91	14.4	16	1	US-07-959-946-5 Sequence 5, Appli
12	91	14.4	16	1	US-08-333-577-5 Sequence 5, Appli
13	91	14.4	16	5	PCT-US92-08634-5 Sequence 6, Appli
14	87.5	13.9	317	4	US-08-949-155-6 Sequence 46, Appli
15	82.5	13.1	515	2	US-08-705-660-46 Sequence 46, Appli
16	82.5	13.1	515	3	US-08-989-045-46 Sequence 29, Appli
17	79	12.5	220	2	US-08-726-306A-29 Sequence 29, Appli
18	79	12.5	230	1	US-08-118-469A-3 Sequence 3, Appli
19	79	12.5	230	1	US-08-909-119-3 Sequence 3, Appli
20	78	12.5	317	1	US-07-709-949-2 Sequence 2, Appli
21	78	12.4	110	1	US-07-849-389-7 Sequence 7, Appli
22	78	12.4	900	2	US-08-630-822A-62 Sequence 62, Appli
23	78	12.4	900	2	US-09-005-069-62 Sequence 62, Appli
24	77	12.2	1786	4	US-08-973-462-8 Sequence 8, Appli
25	76	12.0	107	1	US-08-182-175A-105 Sequence 105, App
26	76	12.0	107	1	US-08-474-633A-92 Sequence 92, Appli
27	76	12.0	107	5	PCT-US92-06412-105 Sequence 105, App

28	75.5	12.0	203	1	US-08-216-593-5	Sequence 5, Appli
29	75.5	12.0	203	5	PCT-US93-12380-5	Sequence 5, Appli
30	75.5	12.0	205	1	US-07-992-827D-5	Sequence 5, Appli
31	75.5	12.0	393	2	US-08-904-031-3	Sequence 3, Appli
32	74.5	11.8	897	1	US-08-095-737-4	Sequence 4, Appli
33	74.5	11.8	897	1	US-08-480-145-4	Sequence 4, Appli
34	74.5	11.8	897	2	US-08-477-389-4	Sequence 4, Appli
35	73.5	11.6	386	4	US-09-085-199B-2	Sequence 2, Appli
36	73.5	11.6	896	1	US-08-095-737-2	Sequence 2, Appli
37	73.5	11.6	896	1	US-08-480-145-2	Sequence 2, Appli
38	73.5	11.6	896	2	US-08-477-389-2	Sequence 2, Appli
39	73.5	11.6	914	4	US-09-085-199B-4	Sequence 4, Appli
40	73.5	11.6	1090	4	US-09-085-199B-5	Sequence 5, Appli
41	73.5	11.6	1093	4	US-09-315-793-52	Sequence 52, Appli
42	73	11.6	707	2	US-08-949-941B-2	Sequence 2, Appli
43	72.5	11.5	392	2	US-08-904-031-1	Sequence 1, Appli
44	72.5	11.5	984	1	US-08-242-932-2	Sequence 2, Appli
45	72.5	11.5	984	1	US-08-714-481-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-448-606-6  
; Sequence 6, Application US/08448606  
; Patent No. 5721114  
; GENERAL INFORMATION:  
; APPLICANT: Abrahams n, Lars  
; APPLICANT: Holmgren, Erik  
; APPLICANT: Kalder n, Christina  
; APPLICANT: Lake, Mats  
; APPLICANT: Mikaelsson, sa  
; APPLICANT: Sejlitz, Torsten  
; TITLE OF INVENTION: Expression System For Producing  
; TITLE OF INVENTION: Apolipoprotein AI-M  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pollock, Vande Sande & Priddy  
; STREET: 1990 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,606  
; FILING DATE: 25-AUG-1995  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/SE93/01061  
; FILING DATE: 09-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9203753-0  
; FILING DATE: 11-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amernick, Burton A.  
; REGISTRATION NUMBER: 24,852  
; REFERENCE/DOCKET NUMBER: 0151/00121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)331-7111  
; TELEFAX: (202) 293-6229  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-448-606-6

Query Match 100.0%; Score 631; DB 1; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2.7e-54;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
|||||  
DB 22 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 81  
QY 61 REQLGPVTQEFWMDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120  
|||||  
DB 82 REQLGPVTQEFWMDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 141

## RESULT 2

US-07-959-946-3  
; Sequence 3, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959, 946  
; FILING DATE: 19921008  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901, 706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-946-3

Query Match 100.0%; Score 631; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.7e-54;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
|||||  
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
QY 61 REQLGPVTQEFWMDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120

DB 85 REQLGPVTQEFWMDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 144  
|||||

## RESULT 3

US-08-333-577-3  
; Sequence 3, Application US/08333577  
; Patent No. 5786206  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5786206th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333, 577  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER: SCRF 234.0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-333-577-3

Query Match 100.0%; Score 631; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.7e-54;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
|||||  
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
QY 61 REQLGPVTQEFWMDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120  
|||||  
DB 85 REQLGPVTQEFWMDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 144

## RESULT 4

US-08-952-796-2  
; Sequence 2, Application US/08952796  
; Patent No. 6258596  
; GENERAL INFORMATION:  
; APPLICANT: BENOIT, Patrick  
; APPLICANT: BRUCKERT, Eric  
; APPLICANT: DENEFE, Patrice  
; APPLICANT: DUBERGER, Nicolas  
; APPLICANT: FRUCHART, Jean-Charles

APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehlner Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-2

Query Match 100.0%; Score 631; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.7e-54;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120  
DB 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 144

RESULT 5  
PCT-US92-08634-3  
Sequence 3, Application PC/TUS9208634  
GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Witzum, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 North Stetson, Suite 4700

CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08634  
FILING DATE: 19921009  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamsen, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-08634-3

Query Match 100.0%; Score 631; DB 5; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.7e-54;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120  
DB 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 144

RESULT 6  
US-08-952-796-15  
Sequence 15, Application US/08952796  
Patent No. 6258596  
GENERAL INFORMATION:  
APPLICANT: BENOIT, Patrick  
APPLICANT: BRUCKERT, Eric  
APPLICANT: DENEFE, Patrice  
APPLICANT: DUBERGER, Nicolas  
APPLICANT: FRUCHART, Jean-Charles  
APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/952,796  
;; FILING DATE:  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/06061  
;; FILING DATE: 22-MAY-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO FR96/00747  
;; FILING DATE: 20-MAY-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fehlner Esq., Paul F.  
;; REGISTRATION NUMBER: 35,135  
;; REFERENCE/DOCKET NUMBER: ST95031-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 200 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-952-796-15

Query Match 64.0%; Score 404; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2.9e-32;  
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 LKLDNDVSTSTFSKLRQLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDD 103  
Db 1 LKLDNDVSTSTFSKLRQLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDD 60

OY 104 FOKKWOEMELLYRQKVE 120  
Db 61 FOKKWOEMELLYRQKVE 77

RESULT 7  
US-08-292-870-2  
; Sequence 2, Application US/08292870  
; Patent No. 5814467  
; GENERAL INFORMATION:  
; APPLICANT: Curtiss, Linda K  
; APPLICANT: Banka, Carole L  
; APPLICANT: Bonnet, David J  
; APPLICANT: Smith, Richard S  
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/292,870  
; FILING DATE: 17-AUG-1994  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/534,761  
; FILING DATE: 07-JUN-1990  
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/711,333  
;; FILING DATE: 06-JUN-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US 91/04038  
;; FILING DATE: 07-JUN-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 619-554-2937  
;; TELEFAX: 619-554-6312  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 44 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; US-08-292-870-2

Query Match 31.4%; Score 198; DB 2; Length 44;  
Best Local Similarity 100.0%; Pred. No. 5.7e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 ROEMSKDLEEVKAKVQPYLDDFQKKWOEMELLYRQKVE 120  
Db 1 ROEMSKDLEEVKAKVQPYLDDFQKKWOEMELLYRQKVE 38

RESULT 8  
US-08-292-870-1  
; Sequence 1, Application US/08292870  
; Patent No. 5814467  
; GENERAL INFORMATION:  
; APPLICANT: Curtiss, Linda K  
; APPLICANT: Banka, Carole L  
; APPLICANT: Bonnet, David J  
; APPLICANT: Smith, Richard S  
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/292,870  
; FILING DATE: 17-AUG-1994  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/534,761  
; FILING DATE: 07-JUN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/711,333  
; FILING DATE: 06-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US 91/04038  
; FILING DATE: 07-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-292-870-1

Query Match 29.8%; Score 188; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 8,7e-12;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 EMSKDLEEVKAKVQPYLDDFQKKWQEMELLYRQKVE 120  
Db 1 EMSKDLEEVKAKVQPYLDDFQKKWQEMELLYRQKVE 36

RESULT 9  
US-08-292-870-3  
Sequence 3, Application US/08292870  
Patent No. 5814467  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
FRAGMENT TYPE: internal

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11  
OTHER INFORMATION: /note="Xaa can be either E (Glu)  
OTHER INFORMATION: or F (Phe)"  
US-08-292-870-3

Query Match 21.9%; Score 138; DB 2; Length 32;  
Best Local Similarity 96.3%; Pred. No. 2.8e-07;  
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 KAKVQPYLDDFQKKWQEMELLYRQKVE 120  
Db 1 KAKVQPYLDDXQKKWQEMELLYRQKVE 27

RESULT 10  
US-08-292-870-4  
Sequence 4, Application US/08292870  
Patent No. 5814467  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal



US-08-292-870-4

Query Match 19.7%; Score 124; DB 2; Length 25;  
Best Local Similarity 96.0%; Pred. No. 4.8e-06;  
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 87 SKDLEEVKAKVQPYLDDEFOKKWQEE 111  
Db 1 SKDLEEVKAKVQPYLDDEFOKKWQEE 25

RESULT 11

US-07-959-946-5

; Sequence 5, Application US/07959946

; Patent No. 5408038

; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.

; APPLICANT: Koduri, Raju

; APPLICANT: Young, Stephen G.

; APPLICANT: Witzum, Joseph L.

; APPLICANT: Curtiss, Linda K.

; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a

; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &amp;

; ADDRESSEE: Milnamow, Ltd.

; STREET: 180 No. 5408038th Stetson, Suite 4700

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/959,946

; FILING DATE: 19921008

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/901,706

; FILING DATE: 18-JUN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Gansson, Edward P.

; REGISTRATION NUMBER: 29,381

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312)616-5400

; TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 16 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-07-959-946-5

Query Match 14.4%; Score 91; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KVQPYLDDEFOKKWQEE 111  
Db 1 KVQPYLDDEFOKKWQEE 16

RESULT 12

US-08-333-577-5

; Sequence 5, Application US/08333577

; Patent No. 5786206

; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.

; APPLICANT: Koduri, Raju

; APPLICANT: Young, Stephen G.

; APPLICANT: Witzum, Joseph L.

; APPLICANT: Curtiss, Linda K.

; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a

; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &amp;

; ADDRESSEE: Milnamow, Ltd.

; STREET: 180 No. 5786206th Stetson, Suite 4700

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/333,577

; FILING DATE:

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Gansson, Edward P.

; REGISTRATION NUMBER: 29,381

; REFERENCE/DOCKET NUMBER: SCRF 234.0

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312)616-5400

; TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 16 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-333-577-5

Query Match 14.4%; Score 91; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KVQPYLDDEFOKKWQEE 111  
Db 1 KVQPYLDDEFOKKWQEE 16

RESULT 13

PCT-US92-08634-5

; Sequence 5, Application PC/TUS9208634

; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.

; APPLICANT: Koduri, Raju

; APPLICANT: Young, Stephen G.

; APPLICANT: Witzum, Joseph L.

; APPLICANT: Curtiss, Linda K.

; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a

; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &amp;

; ADDRESSEE: Milnamow, Ltd.

; STREET: 180 North Stetson, Suite 4700

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60601

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08634  
FILING DATE: 19921009  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamsen, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-08634-5

Query Match 14.4%; Score 91; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KVGPLYDDFQKKWQEE 111  
Db 1 KVGPLYDDFQKKWQEE 16

RESULT 14  
US-08-949-155-6  
Sequence 6, Application US/08949155  
Patent No. 6271436  
GENERAL INFORMATION:  
APPLICANT: Piedrahita, Jorge A  
APPLICANT: Bazer, Fuller W  
TITLE OF INVENTION: Compositions and Methods for the  
TITLE OF INVENTION: Generation of Transgenic Animal Species  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE AND DURKEE  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: US  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/949,155  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/027,338  
FILING DATE: 11-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,094  
FILING DATE: 09-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hilder, David W.  
REGISTRATION NUMBER: 41,071  
REFERENCE/DOCKET NUMBER: TAMK:177  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 317 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-949-155-6

Query Match 13.9%; Score 87.5; DB 4; Length 317;  
Best Local Similarity 20.5%; Pred. No. 0.42;  
Matches 25; Conservative 28; Mismatches 66; Indels 3; Gaps 1;

QY 1 DEPP---QSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTF 57  
Db 33 EEPKWSQSPWEQALGFRWDYLRWVQSLSDQVEELLSTKVQELTELIEESMKEVAYR 92  
QY 58 SKLRQGLPVTQEFWDNLEKETEGLRQENSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQ 117  
Db 93 EELEAQGLPVTQETQARLSKELQAAQARVADMEDVARNRLVLYRSEVHNLGQTTEELRS 152  
QY 118 KV 119  
Db 153 RL 154

RESULT 15  
US-08-705-660-46  
Sequence 46, Application US/08705660  
Patent No. 5858683  
GENERAL INFORMATION:  
APPLICANT: KEESEE, SUSAN  
APPLICANT: OBAR, ROBERT  
APPLICANT: WU, YING-JYE  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
TITLE OF INVENTION: DETECTION OF CERVICAL CANCER  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Testa, Hurwitz & Thibault  
STREET: 125 High St.  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,660  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: GREENHALGH, DUNCAN A  
REGISTRATION NUMBER: 38,678  
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7100  
TELEFAX: (617) 248-7000  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 515 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-705-660-46

Query Match 13.1%; Score 82.5; DB 2; Length 515;



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:05:15 ; Search time 139.03 Seconds  
(without alignments)  
95.870 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_25\_144  
Perfect score: 631  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....LDDEFQKKQWQEMELRYQKVE 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:\*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	631	100.0	243	9	AAp81082	Sequence of mature
2	631	100.0	264	15	AAp56863	Apo-lipoprotein AI
3	631	100.0	264	15	AAp56864	Apo-lipoprotein AI
4	631	100.0	267	7	AAp61079	Assumed human apol
5	631	100.0	267	9	AAp82128	Entire human prepr
6	631	100.0	267	14	AAp34032	Sequence of apo AI
7	631	100.0	267	16	AAp72705	Human apo A-I incl
8	631	100.0	267	18	AAp08602	Human apolipoprote
9	631	100.0	267	20	AAp18675	Human apolipoprote
10	631	100.0	267	22	AAp47620	Full length Apo-AI
11	631	100.0	268	9	AAp80668	Recombinant human

12	631	100.0	299	22	AAU33170	Novel human secret
13	500	79.2	154	22	AAO12095	Human polypeptide
14	489	77.5	151	22	AAO02278	Human polypeptide
15	485.5	76.9	166	22	AAU28372	Novel human secret
16	485.5	76.9	244	22	AAU28184	Novel human secret
17	476.5	75.5	120	22	AAU30267	Novel human secret
18	476.5	75.5	120	22	AAU30469	Novel human secret
19	458	72.6	221	22	AAU29835	Novel human secret
20	410	65.0	119	22	AAU30468	Novel human secret
21	381.5	60.5	318	22	AAU30268	Novel human secret
22	247.5	39.2	359	22	AAU30470	Novel human secret
23	193	30.6	42	13	AAp20164	Apo AI polypeptide
24	142.5	22.6	377	14	AAp39502	Human apoAIV mutei
25	139	22.0	183	14	AAp39482	Human apoAIV mutei
26	139	22.0	333	14	AAp39481	Human apoAIV mutei
27	139	22.0	333	14	AAp39488	Human apoAIV mutei
28	139	22.0	333	14	AAp39490	Human apoAIV mutei
29	139	22.0	333	14	AAp39495	Human apoAIV mutei
30	139	22.0	333	14	AAp39497	Human apoAIV mutei
31	139	22.0	337	14	AAp39492	Human apoAIV mutei
32	139	22.0	342	14	AAp39487	Human apoAIV mutei
33	139	22.0	342	14	AAp39489	Human apoAIV mutei
34	139	22.0	342	14	AAp39491	Human apoAIV mutei
35	139	22.0	342	14	AAp39496	Human apoAIV mutei
36	139	22.0	342	14	AAp39498	Human apoAIV mutei
37	139	22.0	346	14	AAp39493	Human apoAIV mutei
38	139	22.0	377	14	AAp39443	Human apolipoprote
39	139	22.0	377	14	AAp45242	Human apoAIV mutei
40	139	22.0	377	14	AAp45243	Human apoAIV mutei
41	139	22.0	377	14	AAp45244	Human apoAIV mutei
42	139	22.0	396	22	AAp90664	Human secreted pro
43	138	21.9	32	13	AAp20165	Apo AI polypeptide
44	138	21.9	377	14	AAp39501	Human apoAIV mutei
45	137	21.7	337	14	AAp39494	Human apoAIV mutei

## ALIGNMENTS

RESULT 1

AAp81082

AAp81082 standard; protein; 243 AA.

XX

AC AAp81082;

XX

DT 14-JAN-1991 (first entry)

XX

DE Sequence of mature human apolipoprotein AI (apoAI).

XX

KW Atherosclerosis; therapy; cardiovascular disease.

XX

OS Homo sapiens.

XX

PN W08803166-A.

XX

PD 05-MAY-1988.

XX

PE 21-OCT-1987; 87WO-EP00621.

XX

PR 23-OCT-1986; 86GB-0025435.

XX

PA (FARM ) FARMITALIA C ERBA SPA.

XX

PI Lorenzetti R, Monaco L, Soria M, Palomba R, Isacchi A, Sarmientos P;

XX

DR WPI; 1988-133240/19.

XX

DR N-PSDB; AAN80243.

XX

PT Recombinant human apo:lipoprotein AI -

PT used to lower plasma cholesterol and/or tri glyceride levels and

PT to combat atherosclerosis and cardiovascular diseases

XX

PS Disclosure; Fig 1; 51pp; English.

XX The protein comprising apo AI genetic variants may be used to lower  
CC plasma cholesterol and/or triglyceride levels. They may also be used to  
CC combat atherosclerosis and cardiovascular diseases such as coronary  
CC heart disease. Prefd. proteins are Met-apo AI, Met-apo AI-T6, Met-apo  
CC AI-MI and Met-apo AI-T6/MI.  
XX  
SQ Sequence 243 AA;

Query Match 100.0%; Score 631; DB 9; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.2e-50;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 1 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfskl 60  
QY 61 REQLGPVTQEFWMDNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELRYQKVE 120  
Db 61 reqlgpvtqefwmdneketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve 120

## RESULT 2

AAR56863  
ID AAR56863 standard; Protein; 264 AA.

AC AAR56863;

DT 26-JAN-1995 (first entry)

DE Apo-lipoprotein AI-M.

KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid PKP683;  
vector; atherosclerosis; cardiovascular disease.

OS Homo sapiens.

PN WO9413819-A.

PD 23-JUN-1994.

PF 09-DEC-1993; 93WO-SE01061.

PR 11-DEC-1992; 92SE-0003753.

PA (KABI ) KABI PHARMACIA AB.

PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
PI Sejlitz T;

DR WPI; 1994-217892/26.  
DR N-PSDB; AAQ68357.

PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 3; 33pp; English.

CC Plasmid PKP683 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of PKP683 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.

SQ Sequence 264 AA;

Query Match 100.0%; Score 631; DB 15; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.3e-50;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 22 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfskl 81  
QY 61 REQLGPVTQEFWMDNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELRYQKVE 120  
Db 82 reqlgpvtqefwmdneketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve 141

## RESULT 3

AAR56864  
ID AAR56864 standard; Protein; 264 AA.

AC AAR56864;

DT 26-JAN-1995 (first entry)

DE Apo-lipoprotein AI-M.

KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid PKP764;  
vector; atherosclerosis; cardiovascular disease.

OS Homo sapiens.

PN WO9413819-A.

PD 23-JUN-1994.

PF 09-DEC-1993; 93WO-SE01061.

PR 11-DEC-1992; 92SE-0003753.

PA (KABI ) KABI PHARMACIA AB.

PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
PI Sejlitz T;

DR WPI; 1994-217892/26.  
DR N-PSDB; AAQ68358.

PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 4; 33pp; English.

CC Plasmid PKP764 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of PKP764 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.

SQ Sequence 264 AA;

Query Match 100.0%; Score 631; DB 15; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.3e-50;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 22 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfskl 81  
QY 61 REQLGPVTQEFWMDNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELRYQKVE 120  
Db 82 reqlgpvtqefwmdneketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve 141

## RESULT 4

AAP61079  
ID AAP61079 standard; Protein; 267 AA.

AC AAP61079;



XX 07-OCT-1991 (first entry)  
XX Assumed human apolipoprotein A-1 derivative gene product.  
DE Hyperlipaemia; arteriosclerosis.  
XX Homo sapiens.  
XX JP61096998-A.  
XX 15-MAY-1986.  
PD 16-OCT-1984; 84JP-0216988.  
XX 16-OCT-1984; 84JP-0216988.  
XX 16-OCT-1984; 84JP-0216988.  
XX (MITU ) MITSUBISHI CHEM IND KK.  
XX WPI; 1986-165025/26.  
DR N-PSDB; AAN60886.  
XX  
XX Human apo:lipoprotein A-1 (deriv.) prepn. - by providing DNA  
PT fragment in cloning site downstream of expression vector promoter  
PT and introducing info host microorganism.  
XX Disclosure; Fig 2; 9pp; Japanese.  
XX The human apolipoprotein may be produced by a suitable transformed  
CC host, it is effective in treating hyperlipaemia and arteriosclerosis.  
XX  
SQ Sequence 267 AA;  
  
Query Match 100.0%; Score 631; DB 7; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.4e-50;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfskl 84  
  
QY 61 REQLGPTYQEFWDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120  
Db 85 reqlgptvqefwdlneketeglrqemskdleevkakvgpylddfqkkwqeemelyrqkve 144  
  
RESULT 5  
AAP82128  
ID AAP82128 standard; protein; 267 AA.  
XX  
AC AAP82128;  
XX  
DT 24-OCT-1990 (first entry)  
XX  
DE Entire human preproapoprotein A1.  
XX human preproapoprotein A1; high density lipoprotein deficiency; ss.  
XX synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT /label=precursor  
FT Peptide 19..24  
FT /label=propeptide  
FT Protein 25..267  
FT /label-mature apoprotein  
XX  
XX EP293357-A.  
XX  
XX 30-NOV-1988.  
XX

PF 24-MAY-1988; 88EP-0870095.  
XX  
XX 28-MAY-1987; 87GB-0012540.  
XX  
XX (UNIO ) UCB SA.  
XX  
PI Bollen A, Gobert J, Wulfert E;  
XX  
DR WPI; 1988-339891/48.  
DR N-PSDB; AAN82064.  
XX  
XX New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures  
PT  
XX Disclosure; p; French.  
XX  
XX The cDNA 878bp fragment encoding preproapoprotein A1 was detected  
CC in clone PULB1609 derived from human liver cells.  
CC See also AAN81258.  
XX  
SQ Sequence 267 AA;  
  
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfskl 84  
  
QY 61 REQLGPTYQEFWDLNEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 120  
Db 85 reqlgptvqefwdlneketeglrqemskdleevkakvgpylddfqkkwqeemelyrqkve 144  
  
RESULT 6  
AAR34032  
ID AAR34032 standard; Protein; 267 AA.  
XX  
AC AAR34032;  
XX  
DT 13-AUG-1993 (first entry)  
XX  
DE Sequence of apo A1.  
XX  
XX Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.  
XX  
XX Homo sapiens.  
XX  
OS  
XX  
PN WO9307165-A.  
XX  
PD 15-APR-1993.  
XX  
PF 09-OCT-1992; 92WO-US08634.  
XX  
XX 09-OCT-1991; 91US-0774633.  
PR 08-OCT-1992; 92US-055555.  
PR 28-JUN-1992; 92US-0901706.  
XX  
XX (SCRI ) SCRIPPS RES INST.  
XX  
XX  
PI Curtiss LK, Koduri KR, Smith RS, Witztum JL, Young SG;  
XX  
XX WPI; 1993-134378/16.  
DR N-PSDB; AAQ40030.  
XX  
XX Polypeptide mimic of native apo B-100 and native apo A-I - useful  
PT in assays for LDL and HDL in plasma samples  
XX  
XX Claim 19; Pages 105-106; 137pp; English.  
XX  
XX The inventors claim a portion of the polypeptide contg. apo B-100





FT	Domain	90..111
FT	/note= "Involved in lectin-cholesterol acyltransferase activity"	
FT	Domain	44..65
FT	/label= Amphipathic helix	
FT	Domain	66..98
FT	/label= Amphipathic helix	
FT	Domain	99..120
FT	/label= Amphipathic helix	
FT	Domain	121..142
FT	/label= Amphipathic helix	
FT	Domain	143..164
FT	/label= Amphipathic helix	
FT	Domain	165..208
FT	/label= Amphipathic helix	
FT	Domain	209..219
FT	/label= Amphipathic helix	
FT	Domain	220..241
FT	/label= Amphipathic helix	
FT	Peptide	25..194
FT	/label= AFTI	
FT	/note= "18 kD N-terminal fragment"	
FT	Peptide	25..144
FT	/label= AFTI	
FT	/note= "13 kD N-terminal fragment"	
FT	Peptide	156..267
FT	/label= AFTI	
FT	/note= "13 kD C-terminal fragment"	
PN	WO200168852-A2.	
XX		
PD	20-SEP-2001.	
XX		
PX	13-MAR-2001; 2001WO-US07826.	
XX		
PR	13-MAR-2000; 2000US-189008P.	
PA	(AMGE-) AMGEN INC.	
PI	Edwards CK, Burger D, Dayer J, Kohno T;	
XX		
DR	WPI: 2001-596908/67.	
XX	N-PSDB; AAH43623.	
PT	Apo-A-I fragment T-cell activation inhibitor (AFTI) polynucleotides,	
PT	useful for treating, diagnosing, ameliorating diseases associated with	
PT	IL-1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease	
XX	and asthma -	
PS	Claim 1; Fig 1A; 132pp; English.	
XX		
CC	This sequence shows full length apolipoprotein (Apo-AI). Fragments	
CC	of Apo-AI may be used as Apo-A-I fragment T-cell activation inhibitors	
CC	(AFTI). These fragments are selected from an 18 kD N-terminal fragment	
CC	(amino acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144)	
CC	and a 13 kD C-terminal fragment (amino acids 156-267). The AFTI	
CC	polypeptides and polynucleotides are useful for regulating T-cell	
CC	mediated activation of monocytes and for treating, diagnosing,	
CC	ameliorating diseases associated with IL-1 and/or TNF activity.	
CC	The diseases are acute pancreatitis, Alzheimer's disease, asthma,	
CC	cancer, fever, inflammatory bowel disease, ischemia, multiple	
CC	sclerosis, osteoporosis, Parkinson's disease or psoriasis. Numerous	
CC	examples of other diseases are given in the specification.	
CC	The AFTI nucleic acids are useful as hybridization probes in diagnostic	
CC	tissue or bodily fluid samples.	
SQ	Sequence 267 AA;	
XX		
Query Match	100.0%; Score 631; DB 22; Length 267;	
Best Local Similarity	100.0%; Pred. No. 1.4e-50;	
Matches 120; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	

QY	1	DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSVTSTSKL	60
Db	25	deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalsgkqlnlkldnwdsvststskl	84
QY	61	REQLGPVTOEFWMDLEKETEGRLQEMSKDLEEVKAKVQPYLDDFOKKWOEEMELIROYE	120
Db	85	reqlgpvtqefwmdleketegrlqemskdleevkavqpylddfqkkwgeemelypqve	144

RESULT 11  
AAP80668

ID AAP80668 standard; protein; 268 AA.

AC AAP80668;

DT 24-OCT-1990 (first entry)

DE Recombinant human preproapoprotein A1.

KW human preproapoprotein A1; high density lipoprotein deficiency; ss.

synthetic.

FH	Key	Location/Qualifiers
----	-----	---------------------

```

--precursor
1:10
/label=precursor
FT

```

FT	label=protein
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

FT	/label=mature apoprotein A1
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

PN EP293357-A.

PD 30-NOV-1988.

PF 24-MAY-1988; 88EP-0870095.

PR 28-MAY-1987; 87GB-0012540.

PA (UNIO ) UCB SA.

PI Bollen A, Gobert J, Wulfert E;

DR WPI; 1988-339891/48.

XX

PT New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures

PS Claim 1; Page 12; 25pp; French.

CC Met at posn 19 is inserted as an extra amino acid c.f. wild-type

CC The DNA fragment used to replace the wild-type sequence encoding

CC corresponding to wild-type amino acids -6, -1, 1, 3, 4, 5, 6, 7,

CC as in the wild-type protein but reduce formation of secondary

CC See also AAN82064.

SQ Sequence 268 AA;

Query Match	100.0%;	Score 631;	DB 9;	Length 268;
Best Local Similarity	100.0%;	Pred. No. 1.4e-50;		
Matches 120; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]



QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEMELYRQKVE 120  
|  
Db 86 reglpgvtqefwvnlketeqlrgemskdleevkavqpylddfqkkwgeemelyrqkve 145

## RESULT 12

AAU33170

ID AAU33170 standard; Protein; 299 AA.

AC AAU33170;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3661.

KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -

PS Claim 20; Page 718; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

SQ Sequence 299 AA;

Query Match 100.0%; Score 631; DB 22; Length 299;  
Best Local Similarity 100.0%; Pred. No. 1.6e-50;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
|  
Db 57 deppqspwdrvkdlatvyvdvldksgrdyvsqfegsalgkqlnlklldnwdsvtstfsl 116  
QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEMELYRQKVE 120  
|  
Db 117 reglpgvtqefwvnlketeqlrgemskdleevkavqpylddfqkkwgeemelyrqkve 176

## RESULT 13

AAO12095

ID AAO12095 standard; Protein; 154 AA.

AC AAO12095;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 25987.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR N-PSDB; AAI92026.

PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

PS Claim 20; SEQ ID NO 25987; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and other cytokines are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 154 AA;

Query Match 79.2%; Score 500; DB 22; Length 154;  
Best Local Similarity 84.8%; Pred. No. 9.3e-39;  
Matches 95; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
|  
Db 33 deppqspwdrvkdlatvyvdvldksgrdyasqfegyalgkqlnlklldnwdsvtstfsl 92  
QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEMEL 112  
|  
Db 93 reglpgvtqefwvnlketeqlrgemskdleevkavqpylddfqkwtwqkel 144

RESULT 14



AAO02278	ID	AAO02278 standard; Protein; 151 AA.
XX	AC	AAO02278;
XX	DT	06-NOV-2001 (first entry)
XX	DE	Human polypeptide SEQ ID NO 16170.
KW	KM	Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation.
OS	OS	Homo sapiens.
PN	PN	WO200164835-A2.
PD	PD	07-SEP-2001.
PF	PF	26-FEB-2001; 2001WO-US04927.
PR	PR	28-FEB-2000; 2000US-0515126. 18-MAY-2000; 2000US-0577409.
PA	PA	(HYSE-) HYSEQ INC.
PI	PI	Tang YT, Liu C, Drmanac RT;
DR	DR	WPI; 2001-514838/56. N-PSDB; AAI82209.
PT	PT	Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune disorders -
PS	PS	Claim 20; SEQ ID NO 16170; 1399pp + Sequence Listing; English.
CC	CC	The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.
CC	CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SO	SO	Sequence 151 AA;
QY	QY	Query Match 77.5%; Score 489; DB 22; Length 151; Best Local Similarity 94.9%; Pred. No. 9,5e-38; Matches 94; Conservative 1; Mismatches 4; Indels 0; Gaps 0,
DG	DG	1 LKDSGRDVSQFEGSALGKLNKLNDNWDSVTSTFSKLREQLGPVTQEFWNLEKELEG 81   1 LKSGRDYVSQFGCALGKLNKLNDNWDSETSTFSKLREQIGPVTGEFWANLEKDEG 60 
DG	DG	82 LRQMSKDLEEVKAKVQPYLDDFOKKWOEMELRYRQVE 120   61 LRqmskdleeeekakvgpylddfgkkwqeemelyrqkee 99 
RESULT 13		
AAU28372		
AAU28372 standard; Protein; 166 AA.		

AC	AAU28372;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Novel human secretory protein, Seq ID No 729.
XX	
KW	Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW	ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW	transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW	amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW	ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW	gut protection; lung; liver fibrosis; immune deficiency; infection;
KW	severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW	multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW	fertility; analgesic; pain; antigen.
XX	
OS	Homo sapiens.
XX	
PN	WO200166689-A2.
XX	
PD	13-SEP-2001.
XX	
PF	05-MAR-2001; 2001WO-US04942.
XX	
PR	07-MAR-2000; 2000US-0519705.
PR	19-MAY-2000; 2000US-0574454.
PR	17-JUN-2000; 2000US-0596193.
PR	14-JUL-2000; 2000US-0616847.
PR	19-SEP-2000; 2000US-0665363.
PR	20-OCT-2000; 2000US-0693267.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI	Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX	
DR	WPI; 2001-589934/66.
DR	N-PSDB; AAS45272.
XX	
PT	Novel polypeptides and nucleic acids obtained from cDNA libraries
PT	prepared from various human tissues, for diagnosis and treatment of
PT	cancer, neurological, inflammatory, and autoimmune disorders -
XX	
PS	Example 2; SEQ ID No 729; 107pp; English.
XX	
CC	The invention relates to novel isolated human secreted polypeptides (I)
CC	and polynucleotides (II). (I) and (II) are useful for treating
CC	inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC	ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC	involved in increasing haematopoiesis, stem cell survival, bone growth
CC	and remodeling. (I), (II) and modulators of (II) are useful for
CC	prophylaxis or treatment of one or more cancers. (II) is also useful for
CC	creating transgenic animals useful for studying modulators of the polypeptides.
CC	(I) induces the proliferation of neural cells and regeneration of nerve
CC	and brain tissue and is useful for the treatment of central and
CC	peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC	Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC	sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC	activity, regulation of haematopoiesis and is useful for treating myeloid
CC	or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC	and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC	tissue growth, and in tissue repair, healing of burns, incisions,
CC	ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC	disorders, or periodontal disease. Furthermore, (I) is also useful for
CC	gut protection or regeneration and treatment of lung or liver fibrosis,
CC	reperfusion injury in various tissues, various immune deficiencies and
CC	disorders including severe combined immunodeficiency (SCID), bacterial or
CC	fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC	rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC	reactions and conditions, such as asthma or other respiratory problems.
CC	In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC	fertility, metabolism, catabolism, anabolism, storage or elimination of

CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
 CC analgesic effects or other pain reducing effects, immunoglobulin like  
 CC activity and can act as an antigen in a vaccine composition to raise an  
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein  
 CC amino acid sequences of the invention.  
 XX

SQ Sequence 166 AA;

Query Match 76.98; Score 485.5; DB 22; Length 166;  
 Best Local Similarity 80.0%; Pred. No. 2.2e-37;  
 Matches 96; Conservative 1; Mismatches 0; Indels 23; Gaps 1;

QY	1	DEPPQSPWDRVKDIALTVYVDYLVKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL	60
Db	34	deppqspwdrvkdiatvyvdvdkdsgk-----dsvtstfskl	70
QY	61	REQLGPVTQEFWNLKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE	120
Db	71	reqlgpvtqefwlnleketeglrqemskdleevkakvqpylddfqkkwqeemelyrqkve	130

Search completed: September 22, 2002, 12:05:15  
 Job time: 290 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:22:55 ; Search time 26.49 Seconds  
(without alignments)  
1743.663 Million cell updates/sec

Title: US-09-803-918A-2  
Perfect score: 267  
Sequence: 1 MKAVALTLAVLFLTGSOARR.....SFRVSLSLALEEYTKLNTQ 267

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	12.7	67	4 Q9Y355	Q9Y355 homo sapien
2	26	9.7	26	4 Q9UCT8	Q9UCT8 homo sapien
3	19	7.1	56	6 O02762	O02762 ovls aries
4	19	7.1	264	11 Q9Z2L4	Q9Z2L4 mesocricetu
5	13	4.9	82	6 Q29248	Q29248 sus scrofa
6	12	4.5	258	11 O09054	O09054 rattus norv
7	12	4.5	258	11 O08877	O08877 rattus norv
8	12	4.5	263	11 O09042	O09042 mus musculu
9	12	4.5	263	11 O08855	O08855 mus musculu
10	10	3.7	132	16 Q9HWR9	Q9HWR9 pseudomonas
11	8	3.0	69	11 Q9JCV9	Q9JCV9 mus musculu
12	8	3.0	72	5 Q9UAA7	Q9UAA7 conus livid
13	8	3.0	300	16 Q9HW34	Q9HW34 pseudomonas
14	8	3.0	387	16 Q9PDM0	Q9PDM0 xylella fas
15	8	3.0	395	11 Q9DBN0	Q9DBN0 mus musculu
16	8	3.0	395	11 Q91XF8	Q91XF8 mus musculu

17	8	3.0	413	15 O11556	O11556 human immun
18	8	3.0	435	11 Q01488	Q01488 mus musculu
19	8	3.0	691	5 Q9VZE7	Q9VZE7 drosophila
20	8	3.0	714	2 Q9Z3F4	Q9Z3F4 xanthomonas
21	8	3.0	770	5 Q9GPD3	Q9GPD3 drosophila
22	8	3.0	771	5 Q9GPD5	Q9GPD5 drosophila
23	8	3.0	772	5 Q9GPD4	Q9GPD4 drosophila
24	8	3.0	773	5 Q9GPD8	Q9GPD8 drosophila
25	8	3.0	3562	2 Q9F829	Q9F829 micromonosop
26	8	3.0	5060	2 Q9K5M1	Q9K5M1 anabaena sp
27	7	2.6	9	4 Q9UCS8	Q9UCS8 homo sapien
28	7	2.6	41	4 Q13587	Q13587 homo sapien
29	7	2.6	63	15 Q37693	Q37693 human immun
30	7	2.6	67	10 Q81764	Q81764 arabidopsis
31	7	2.6	72	4 Q14970	Q14970 homo sapien
32	7	2.6	73	4 Q75702	Q75702 homo sapien
33	7	2.6	78	12 Q9DKM9	Q9DKM9 spodoptera
34	7	2.6	83	10 Q942Y3	Q942Y3 oryza sativ
35	7	2.6	87	15 Q9DXP2	Q9DXP2 human immun
36	7	2.6	88	15 Q9EAT6	Q9EAT6 human immun
37	7	2.6	88	15 Q9EAM4	Q9EAM4 human immun
38	7	2.6	88	15 Q90RW4	Q90RW4 human immun
39	7	2.6	89	4 Q9BR82	Q9BR82 homo sapien
40	7	2.6	89	15 Q9ID03	Q9ID03 human immun
41	7	2.6	90	15 Q78841	Q78841 human immun
42	7	2.6	91	15 Q9W9A0	Q9W9A0 human immun
43	7	2.6	91	15 Q9DK41	Q9DK41 human immun
44	7	2.6	92	15 Q9DK65	Q9DK65 human immun
45	7	2.6	92	15 Q77305	Q77305 human immun
46	7	2.6	94	15 Q9WNU8	Q9WNU8 human immun
47	7	2.6	95	15 Q91JUS	Q91JUS human immun
48	7	2.6	95	15 Q75278	Q75278 human immun
49	7	2.6	96	15 Q9WL10	Q9WL10 human immun
50	7	2.6	100	15 Q9IDT9	Q9IDT9 human immun
51	7	2.6	101	15 Q72382	Q72382 human immun
52	7	2.6	101	15 Q50067	Q50067 mycobacteri
53	7	2.6	101	16 Q98TJ9	Q98TJ9 platichtys
54	7	2.6	102	13 Q71020	Q71020 human immun
55	7	2.6	113	15 Q71022	Q71022 human immun
56	7	2.6	113	15 Q78358	Q78358 human immun
57	7	2.6	113	15 Q71024	Q71024 human immun
58	7	2.6	114	15 Q74742	Q74742 human immun
59	7	2.6	114	15 Q71017	Q71017 human immun
60	7	2.6	114	15 Q71376	Q71376 human immun
61	7	2.6	120	15 Q9Q7Z7	Q9Q7Z7 human immun
62	7	2.6	121	15 Q03686	Q03686 actinastrium
63	7	2.6	122	8 Q00245	Q00245 homo sapien
64	7	2.6	129	4 Q46206	Q46206 ascaris lum
65	7	2.6	134	5 Q46208	Q46208 ascaris lum
66	7	2.6	134	5 Q9TYJ8	Q9TYJ8 caenorhabdi
67	7	2.6	150	5 Q9CHN5	Q9CHN5 lactococcus
68	7	2.6	150	16 Q9L6P0	Q9L6P0 salmonella
69	7	2.6	153	2 Q9H8K0	Q9H8K0 homo sapien
70	7	2.6	163	4 Q901E4	Q901E4 human immun
71	7	2.6	170	15 Q901E3	Q901E3 human immun
72	7	2.6	170	15 Q9H674	Q9H674 homo sapien
73	7	2.6	181	4 Q901W2	Q901W2 human immun
74	7	2.6	181	4 Q901W2	Q901W2 human immun
75	7	2.6	185	15 Q901W2	Q901W2 human immun

ALIGNMENTS

RESULT 1  
ID Q9Y355 PRELIMINARY; PRT; 67 AA.  
AC Q9Y355;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE APOLIPROTEIN A1 (FRAGMENT).  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99216322; PubMed=10198255;  
RA Hamidi Asl K., Liepnieks J.J., Nakamura M., Parker F., Benson M.D.;  
RT "A novel apolipoprotein A-I variant, Arg173Pro, associated with  
RT cardiac and cutaneous amyloidosis."  
RL Biochem. Biophys. Res. Commun. 257:584-588(1999).  
DR EMBL; AF148963; AAD34604.1; -.  
DR HSSP; P02647; LAV1.  
KW Lipoprotein.  
FT NON\_TER 1 1  
SQ SEQUENCE 67 AA; 7433 MW; 525E1FEB7BDD5AFB CRC64;

Query Match 12.7%; Score 34; DB 4; Length 67;  
Best Local Similarity 100.0%; Pred. No. 5e-26;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 LAARLEALKENGARLAEYHAKATEHLSTLSEKA 231  
DB 34 LAARLEALKENGARLAEYHAKATEHLSTLSEKA 67

RESULT 2  
O9UCT8 PRELIMINARY; PRT; 26 AA.  
AC O9UCT8;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE SPERM ACTIVATING PROTEIN SUBUNIT I, APOLIPOPROTEIN A1, SPAP SUBUNIT I.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=91369902; PubMed=1909888;  
RA Akerlof E., Jorvall H., Slotte H., Pousette A.;  
RT "Identification of apolipoprotein A1 and immunoglobulin as components  
RT of a serum complex that mediates activation of human sperm motility."  
RL Biochemistry 30:8986-8990(1991).  
SQ SEQUENCE 26 AA; 2930 MW; 64921A333E768D27 CRC64;

Query Match 9.7%; Score 26; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DEPPQSPWDRVKDLATVYVDVLDKDSG 50  
DB 1 DEPPQSPWDRVKDLATVYVDVLDKDSG 26

RESULT 3  
O02762 PRELIMINARY; PRT; 56 AA.  
AC O02762;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOLIPOPROTEIN A1 (FRAGMENT).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=99098454; PubMed=9883985;  
RA Robertson J.A., Bhattacharya S., Ing N.H.;  
RT "Tamoxifen up-regulates oestrogen receptor-alpha, c-fos and  
RT glyceraldehyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium."  
DR EMBL; U94720; AAB57840.1; -.  
DR HSSP; P02647; LAV1.  
KW Lipoprotein.  
FT NON\_TER 1 1  
SQ SEQUENCE 56 AA; 6617 MW; 2AB38E08F1E8F1BC CRC64;

Query Match 7.1%; Score 19; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 4.1e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 REQLGPVTQEFWDNLEKET 103  
DB 15 REQLGPVTQEFWDNLEKET 33

RESULT 4  
O9Z2L4 PRELIMINARY; PRT; 264 AA.  
AC O9Z2L4;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE APOLIPOPROTEIN A-I.  
GN APOAI.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;  
RX MEDLINE=99061559; PubMed=9843713;  
RA Wu J.Y.J., Reaves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;  
RT "Zinc deficiency decreases plasma level and hepatic mRNA abundance of  
RT apolipoprotein A-I in rats and hamsters."  
RL Am. J. Physiol. 275:C1516-C1525(1998).  
DR EMBL; AF046919; AAC98484.1; -.  
DR HSSP; P02647; LAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;

Query Match 7.1%; Score 19; DB 11; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EQLGPVTQEFWDNLEKETE 104  
DB 85 EQLGPVTQEFWDNLEKETE 103

RESULT 5  
O29248 PRELIMINARY; PRT; 82 AA.  
AC O29248;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE APOLIPOPROTEIN A-I (FRAGMENT).  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9623;



RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SMALL INTESTINE;  
RX MEDLINE=96327607; PubMed=8672129;  
RA Winteroe A.K., Fredholm M., Davies W.;  
RT "Evaluation and characterization of a porcine small intestine cDNA  
library."  
RL Mamm. Genome 7:509-517(1996).  
DR EMBL; F14858; CAA23298.1; -.  
KW lipoprotein.  
FT NON\_TER 82 82  
SQ SEQUENCE 82 AA; 9168 MW; 24625C65CBFFEDD8 CRC64;

Query Match 4.9%; Score 13; DB 6; Length 82;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AVLFLTGSQARHF 21  
| | | | | | | | | |  
Db 9 AVLFLTGSQARHF 21

RESULT 6  
ID 009054 PRELIMINARY; PRT; 258 AA.  
AC 009054;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WKY, AND SHRS; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79578; AAB58430.1; -.  
DR EMBL; U79577; AAB58429.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 258 AA; 29918 MW; 093E6EF2E629CDC8 CRC64;

Query Match 4.5%; Score 12; DB 11; Length 258;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KDSGRDYVSQFE 58  
| | | | | | | | | |  
Db 45 KDSGRDYVSQFE 56

RESULT 7  
ID 008877 PRELIMINARY; PRT; 258 AA.  
AC 008877;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SHR; TISSUE=SPLEEN;

RX MEDLINE=98077648; PubMed=9415807;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
gene."  
RL Biochem. Mol. Biol. Int. 43:989-996(1997).  
DR EMBL; U79576; AAB58428.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 258 AA; 29831 MW; 093FB582E629CDC8 CRC64;

Query Match 4.5%; Score 12; DB 11; Length 258;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KDSGRDYVSQFE 58  
| | | | | | | | | |  
Db 45 KDSGRDYVSQFE 56

RESULT 8  
ID 009042 PRELIMINARY; PRT; 263 AA.  
AC 009042;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ICR AND BALB/C, AND C3H; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79575; AAB58427.1; -.  
DR EMBL; U79572; AAB58424.1; -.  
DR EMBL; U79573; AAB58425.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 263 AA; 30516 MW; E245DF7483A5A0DD CRC64;

Query Match 4.5%; Score 12; DB 11; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KDSGRDYVSQFE 58  
| | | | | | | | | |  
Db 46 KDSGRDYVSQFE 57

RESULT 9  
ID 008855 PRELIMINARY; PRT; 263 AA.  
AC 008855;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=SPLEEN;

RX MEDLINE=98077648; PubMed=9415807;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
RT gene."  
RL Biochem. Mol. Biol. Int. 43:989-996(1997).  
DR EMBL; U79574; AAB58426.1; -.  
DR HSSP; P02647; 1AV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 263 AA; 30544 MW; 7825DF7483A5B24A CRC64;

Query Match 4.5%; Score 12; DB 11; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KDSGRDVSQFE 58  
|||||  
DB 46 KDSGRDVSQFE 57

RESULT 10  
Q9HWR9 PRELIMINARY; PRT; 132 AA.  
AC Q9HWR9;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA4111.  
GN PA4111.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL; AE004827; AAC07498.1; -.  
DR InterPro; IPR001109; HupF\_HypC.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 132 AA; 15113 MW; A56B0BD1D9E3EAA0 CRC64;

Query Match 3.7%; Score 10; DB 16; Length 132;  
Best Local Similarity 100.0%; Pred. No. 0.09;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLARLEAL 205  
|||||  
DB 41 QRLARLEAL 50

RESULT 11  
Q99JV9 PRELIMINARY; PRT; 69 AA.  
AC Q99JV9;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 7.9 KDA PROTEIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005635; AAH05635.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 69 AA; 7862 MW; 5996C0B406CA5F4 CRC64;

Query Match 3.0%; Score 8; DB 11; Length 69;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 ALGKQLNL 68  
|||||  
DB 47 ALGKQLNL 54

RESULT 12  
Q9UAA7 PRELIMINARY; PRT; 72 AA.  
AC Q9UAA7;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE FOUR-LOOP CONOTOXIN LVVIA (FRAGMENT).  
OS Conus lividus.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=89426;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RX MEDLINE=99289555; PubMed=10359796;  
RA Duda T.F. Jr., Palumbi S.R.;  
RT "Molecular genetics of ecological diversification: duplication and  
RT rapid evolution of toxin genes of the venomous gastropod Conus."  
RL Proc. Natl. Acad. Sci. U.S.A. 96:6820-6823(1999).  
DR EMBL; AF089912; AAD48167.1; -.  
DR InterPro; IPR004214; Conotoxin.  
DR Pfam; PF02950; Conotoxin; 1.  
FT NON\_TER 1  
SQ SEQUENCE 72 AA; 7970 MW; B7A0DDDC87533D8D CRC64;

Query Match 3.0%; Score 8; DB 5; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AVLFLLTGS 16  
|||||  
DB 5 AVLFLLTGS 12

RESULT 13  
Q9HW34 PRELIMINARY; PRT; 300 AA.  
AC Q9HW34;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA4368.  
GN PA4368.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
RT opportunistic pathogen."  
RL Nature 406:959-964 (2000).  
DR EMBL: AE004852; AAG07756.1; -.  
DR InterPro: IPR002048; EF-hand.  
DR PROSITE: PS00018; EF\_HAND; UNKNOWN\_1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 300 AA; 33580 MW; 7D9ED17502137BA7 CRC64;

Query Match 3.0%; Score 8; DB 16; Length 300;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 RLARLEA 204  
|||  
Db 228 RLARLEA 235

RESULT 14  
Q9PDM0 PRELIMINARY; PRT; 387 AA.  
AC Q9PDM0;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE HYPOTHETICAL PROTEIN XF1359.  
GN XF1359.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidans J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."  
RL Nature 406:151-159 (2000).  
DR EMBL: AE003967; AAF84168.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 387 AA; 42872 MW; 82C09EC7EC7B7027 CRC64;

Query Match 3.0%; Score 8; DB 16; Length 387;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAVLTIAV 10  
|||  
Db 161 AAVLTIAV 168

RESULT 15  
Q9DBN0 PRELIMINARY; PRT; 395 AA.  
AC Q9DBN0;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE ADULT MALE LIVER CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
DE CLONE:1300002K10, FULL INSERT SEQUENCE.  
DE APOA4.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=LIVER;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arai K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690 (2001).  
DR EMBL: AK004856; BAB23620.1; -.  
DR MGD; MGI:88051; APOA4.  
DR InterPro: IPR00074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
SQ SEQUENCE 395 AA; 45044 MW; 4102D84ACB0D182A CRC64;

Query Match 3.0%; Score 8; DB 11; Length 395;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KAAVLTIA 9  
|||  
Db 4 KAAVLTIA 11

RESULT 16  
Q91XF8 PRELIMINARY; PRT; 395 AA.  
AC Q91XF8;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE UNKNOWN (PROTEIN FOR MGC:18592).

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC010769; AAH10769.1; -  
SQ SEQUENCE 395 AA; 45029 MW; C48BE32EED441F71 CRC64;

Query Match 3.0%; Score 8; DB 11; Length 395;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAAVLTLA 9  
Db 4 KAAVLTLA 11

RESULT 17  
ID 011556 PRELIMINARY; PRT; 413 AA.  
AC 011556;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Balotta C., Riva C., Violin M., Colombo C., Galli M., Moroni M.;  
RT "Analysis of the VI-V5 env region in Long-Term Non Progressor and in  
RT rapid progressor HIV-1 infected individuals.";  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U95459; AAB52806.1; -  
DR InterPro; IPR000777; GP120.1.  
DR Pfam; PF00516; GP120.1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT 413 413  
SQ SEQUENCE 413 AA; 45471 MW; 893560C04CE12295 CRC64;

Query Match 3.0%; Score 8; DB 15; Length 413;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 SKLREQLG 89  
Db 238 SKLREQLG 245

RESULT 18  
ID 001488 PRELIMINARY; PRT; 435 AA.  
AC 001488;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE VARIANT APOLIPOPROTEIN A-IV PRECURSOR (APOA-IV).  
GN APOA-4.  
OS Mus musculus castaneus (southeastern Asian house mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10091;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=91286309; PubMed=1648102;  
RA Reue K., Leete T.H.;  
RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and  
RT deletion in a region of tandem repeats.";  
RL J. Biol. Chem. 266:12715-12721(1991).  
CC -I- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
CC COMPONENT OF HDL AND CHYLOMICRONS. IT MAY PLAY A ROLE IN THE  
CC INTRAVASCULAR METABOLISM OF HDL. POTENTIAL ROLE IN CELLULAR  
CC CHOLESTEROL EFFLUX.  
CC -I- TISSUE SPECIFICITY: INTESTINE, LIVER, AND PLASMA.  
CC -I- MISCELLANEOUS: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS  
CC (EACH 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A & B, RELATED 11-  
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
CC HELICAL, & MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH  
CC LECITHIN:CHOLESTEROL ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
CC -I- MISCELLANEOUS: THE APOA-IV POLYMORPHISMS ARE CLASSIFIED BASED ON  
CC THE OCCURRENCE OF THREE, FOUR, OR FIVE E-Q-A-Q/ E-Q-Q-Q/ E-Q-V-Q  
CC REPEAT UNITS.  
DR EMBL; M64250; AAA37216.1; -  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipid transport; HDL; VLDL; Chylomicron; Repeat; Signal; Liver;  
KW Plasma; Multigene family; Polymorphism; Cholesterol metabolism.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 435 VARIANT APOLIPOPROTEIN A-IV.  
FT DOMAIN 374 393 TANDEM REPEATS.  
FT REPEAT 374 377 1.  
FT REPEAT 378 381 2.  
FT REPEAT 382 385 3.  
FT REPEAT 386 389 4.  
FT REPEAT 390 393 5.  
SQ SEQUENCE 435 AA; 49254 MW; B2AF55EAE8E456B2 CRC64;

Query Match 3.0%; Score 8; DB 11; Length 435;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAAVLTLA 9  
Db 4 KAAVLTLA 11

RESULT 19  
ID 09VZE7 PRELIMINARY; PRT; 691 AA.  
AC 09VZE7;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE CG1311 PROTEIN.  
GN CG1311.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,



RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foslér C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003480; AAF47876.1;  
DR FlyBase; FBgn0035523; CG1311.  
SQ SEQUENCE 691 AA; 76840 MW; 73309ED51C733208 CRC64;

Query Match 3.0%; Score 8; DB 5; Length 691;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 AVLTAVL 11  
DB 281 AVLTAVL 288

RESULT 20  
Q9Z3F4 PRELIMINARY; PRT; 714 AA.  
AC Q9Z3F4;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE AVIRULENCE PROTEIN AVRBS2.  
GN AVRBS2.  
OS Xanthomonas campestris (pv. vesicatoria).  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xanthomonas.  
OX NCBI\_TaxID=341;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96345635; PubMed=8755898;  
RA Swords K.M., Dahlbeck D., Kearney B., Roy M., Staskawicz B.J.;  
RT "Spontaneous and induced mutations in a single open reading frame  
alter both virulence and avirulence in *Xanthomonas campestris* pv.  
vesicatoria avrBs2,"  
RT vesicatoria avrBs2,"  
RL J. Bacteriol. 178:4661-4669(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Swords K.M., Dahlbeck D., Kearney B., Roy M., Staskawicz B.J.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF114720; AAD11434.1;  
DR InterPro; IPR004129; GDPD.  
DR Pfam; PF03009; GDPD; 1.  
FT VARIANT 231 231 F -> S (IN REF. 1).  
SQ SEQUENCE 714 AA; 78487 MW; 79C034E286FDC145 CRC64;

Query Match 3.0%; Score 8; DB 2; Length 714;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 ALGQNL 68  
DB 180 ALGQNL 187

RESULT 21  
Q9GPD3 PRELIMINARY; PRT; 770 AA.  
AC Q9GPD3;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE ACP36DE (FRAGMENT).  
GN ACP36DE.  
OS *Drosophila simulans* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SIM8;  
RA Begun D., Whitely P., Todd B., Waldrip H., Clark A.;  
RT "Molecular population genetics of male accessory gland proteins in  
RT *Drosophila*,"  
RL Genetics 0:0-0(2001).  
DR EMBL; AY010674; AAG37418.1;  
DR FlyBase; FBgn0041931; Dsim\Acp36DE.  
FT NON\_TER 1 1  
FT NON\_TER 1 1  
SQ SEQUENCE 770 AA; 86139 MW; 640DFA51AB2B340B CRC64;

Query Match 3.0%; Score 8; DB 5; Length 770;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 KQNLKLL 71  
DB 216 KQNLKLL 223

RESULT 22  
Q9GPD5 PRELIMINARY; PRT; 771 AA.  
AC Q9GPD5;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE ACP36DE (FRAGMENT).  
GN ACP36DE.  
OS *Drosophila simulans* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SIM4;  
RA Begun D., Whitely P., Todd B., Waldrip H., Clark A.;  
RT "Molecular population genetics of male accessory gland proteins in  
RT *Drosophila*,"  
RL Genetics 0:0-0(2001).  
DR EMBL; AY010672; AAG37416.1;  
DR FlyBase; FBgn0041931; Dsim\Acp36DE.  
FT NON\_TER 1 1  
FT NON\_TER 1 1  
SQ SEQUENCE 771 AA; 86139 MW; 640DFA51AB2B340B CRC64;



SO SEQUENCE 771 AA; 86183 MW; ED3D1D945898CA16 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 5; Length 771;  
Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 KQNLKLL 71  
|||||  
Db 214 KQNLKLL 221

RESULT 23  
09GPD4

ID 09GPD4 PRELIMINARY; PRT; 772 AA.

AC 09GPD4;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE ACP36DE (FRAGMENT).

GN ACP36DE.

OS Drosophila simulans (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7240;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SIM7;

RA Begun D., Whitley P., Todd B., Waldrup H., Clark A.;

RT "Molecular population genetics of male accessory gland proteins in

RT Drosophila."

RL Genetics 0:0-0(2001).

DR EMBL; AY010673; AAG37417.1; .

FLYBase; FBgn0041931; Dsim\Acp36DE.

FT NON\_TER 1 1

FT NON\_TER 772 772

SO SEQUENCE 772 AA; 86317 MW; 6EA11326DA49B0C7 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 5; Length 772;  
Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 KQNLKLL 71  
|||||  
Db 216 KQNLKLL 223

RESULT 24  
09GPD8

ID 09GPD8 PRELIMINARY; PRT; 773 AA.

AC 09GPD8;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE ACP36DE (FRAGMENT).

GN ACP36DE.

OS Drosophila simulans (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7240;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SIM1;

RA Begun D., Whitley P., Todd B., Waldrup H., Clark A.;

RT "Molecular population genetics of male accessory gland proteins in

RT Drosophila."

RL Genetics 0:0-0(2001).

DR EMBL; AY010669; AAG37413.1; .

FLYBase; FBgn0041931; Dsim\Acp36DE.

FT NON\_TER 1 1

FT NON\_TER 773 773  
SO SEQUENCE 773 AA; 86429 MW; F059CA84C0267A10 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 5; Length 773;  
Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 KQNLKLL 71  
|||||  
Db 216 KQNLKLL 223

RESULT 25  
Q9F829

ID Q9F829 PRELIMINARY; PRT; 3562 AA.

AC Q9F829;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE MEGALOMICIN 6-DEOXYERYTHRONOLIDE B SYNTHASE 2.

GN MEGALIT.

OS Micromonospora megalomicea subsp. nigra.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Micromonosporinae; Micromonosporaceae;

OC Micromonospora.

OX NCBI\_TaxID=136926;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL3275;

RX MEDLINE=20430101; PubMed=10972798;

RA Volchegursky Y., Hu Z., Katz L., McDaniel R.;

RT "Biosynthesis of the anti-parasitic agent megalomicin: transformation

RT of erythromycin to megalomicin in Saccharopolyspora erythraea."

RL Mol. Microbiol. 37:752-762(2000).

DR EMBL; AF263245; AAG13918.1; .

DR HSSP; P25715; 1MA.

DR InterPro; IPR001227; Acyltransf\_domain.

DR InterPro; IPR002085; Adh\_zn\_family.

DR InterPro; IPR000794; Ketoacyl-synt.

DR InterPro; IPR003880; Phosphopant\_attach.

DR Pfam; PF00698; Acyl\_transf; 2.

DR Pfam; PF00107; adh\_zinc; 1.

DR Pfam; PF02801; ketoacyl-synt; 2.

DR PROSITE; PS50075; ACP\_DOMAIN; 2.

DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.

KW Phosphopantetheine.

SO SEQUENCE 3562 AA; 373617 MW; 4810786A14EC540B CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 2; Length 3562;  
Pred. No. 1.9e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 LRQRLAAR 201  
|||||  
Db 1430 LRQRLAAR 1437

RESULT 26  
Q9K5M1

ID Q9K5M1 PRELIMINARY; PRT; 5060 AA.

AC Q9K5M1;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PEPTIDE SYNTHETASE.

GN ADPB.

OS Anabaena sp. 90.

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

OX NCBI\_TaxID=46234;

RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=90;  
RX MEDLINE=20392447; PubMed=10931313;  
RA Rouhainen L., Paulin L., Suomalainen S., Hyttinen H., Buikema W.,  
RA Haselkorn R., Sivonen K.;  
RT "Genes encoding synthetases of cyclic depsipeptides,  
anabaenopeptidides, in Anabaena strain 90.";  
RL EMBL: AJ269505; CAC01604.1; -  
DR HSP; P14687; 1AMU.  
DR InterPro: IPR000873; AMP-bind.  
DR InterPro: IPR001242; DUF4.  
DR InterPro: IPR001601; Meth-transf.  
DR InterPro: IPR003880; Phosphopant\_attach.  
DR InterPro: IPR000051; SAM\_bind.  
DR Pfam: PF00501; AMP-binding; 4.  
DR Pfam: PF00668; Condensation; 4.  
DR PRINTS: PR00154; AMPBINDING.  
DR PROSITE: PS50075; ACP\_DOMAIN; 4.  
DR PROSITE: PS00455; AMP\_BINDING; 4.  
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_2.  
KW Phosphopantetheine.  
SQ SEQUENCE 5060 AA; 565934 MW; 0A6D498ABC6903E CRC64;

Query Match 3.0%; Score 8; DB 2; Length 5060;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 ELRQRLAA 200  
ID 11111111  
DB 4821 ELRQRLAA 4828

RESULT 27  
Q9UCS8  
ID Q9UCS8 PRELIMINARY; PRT; 9 AA.  
AC Q9UCS8;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
DE APOLIPROTEIN A-I.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92075698; PubMed=1742316;  
RA Ehnholm C., Bozas S.E., Tenkanen H., Kirsbaum L., Metso J.,  
RA Murphy B., Walker I.D.;  
RT "The apolipoprotein A-I binding protein of placenta and the SP-40,40  
protein of human blood are different proteins which both bind to  
RT apolipoprotein A-I.";  
RL Biochim. Biophys. Acta 1086:255-260(1991).  
SQ SEQUENCE 9 AA; 981 MW; 7EE37775A6C7776B CRC64;

Query Match 2.6%; Score 7; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 DEPPQSP 31  
ID 11111111  
DB 1 DEPPQSP 7

RESULT 28  
Q13587  
ID Q13587 PRELIMINARY; PRT; 41 AA.  
AC Q13587;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE FATTY ACID SYNTHASE (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=96278786; PubMed=8662758;  
RA Hsu M.H., Chirala S.S., Wakil S.J.;  
RT "Human fatty-acid synthase gene. Evidence for the presence of two  
RT promoters and their functional interaction.";  
RL J. Biol. Chem. 271:13584-13592(1996).  
DR EMBL: U52428; AAC50536.1; -  
FT NON\_TER 41  
SQ SEQUENCE 41 AA; 4668 MW; ECA60F116AB6E729 CRC64;

Query Match 2.6%; Score 7; DB 4; Length 41;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 QEFWDNL 99  
ID 11111111  
DB 20 QEFWDNL 26

RESULT 29  
O37693  
ID O37693 PRELIMINARY; PRT; 63 AA.  
AC O37693;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DKID;  
RX MEDLINE=98080439; PubMed=9420250;  
RA van't Wout A.B., Ran L.J., Kuiken C.L., Kootstra N.A., Pals S.T.,  
RA Schuitemaker H.;  
RT "Analysis of the temporal relationship between human immunodeficiency  
RT virus type 1 quasispecies in sequential blood samples and various  
RT organs obtained at autopsy.";  
RL J. Virol. 72:488-496(1998).  
DR EMBL: AF021606; AAC04060.1; -  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 63  
SQ SEQUENCE 63 AA; 7341 MW; 53991BE85AD01767 CRC64;

Query Match 2.6%; Score 7; DB 15; Length 63;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 KLREQLG 89  
ID 11111111  
DB 52 KLREQLG 58

RESULT 30  
O81764  
ID O81764 PRELIMINARY; PRT; 67 AA.  
AC O81764;  
DT 01-NOV-1998 (Tremblrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, last annotation update)  
DE HYPOTHEtical 7.6 KDA PROTEIN.  
GN F1715.150 OR ATAG33960.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Vitale D., Liguori R., Argirion A., De Simone V., Hoheisel J.,  
RA Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C.,  
RA Bevan M.;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Vitale D., Liguori R., Argirion A., De Simone V., Mewes H.W.,  
RA Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL031032; CAA19878.1; -.  
DR EMBL; AL161584; CAB80113.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 67 AA; 7600 MW; 241F65FFFD865CF CRC64;

Query Match 2.6%; Score 7; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 LRQGLLP 244  
Db 46 LRQGLLP 52

RESULT 31  
Q14970 PRELIMINARY; PRT; 72 AA.  
AC Q14970;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)  
DE INTEGRAL MEMBRANE PROTEIN (FRAGMENT).  
GN NRAMPI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95053705; PubMed=7964458;  
RA Cellier M.F., Govoni G., Vidal S., Kwan T., Groulx N., Liu J.,  
RA Sanchez F., Skamene E., Schurr E., Gros P.;  
RT "Human natural resistance-associated macrophage protein: cDNA cloning,  
RT chromosomal mapping, genomic organization and tissue-specific  
RT expression";  
RL J. Exp. Med. 180:1741-1752(1994).  
DR EMBL; L38593; AAA57556.1; -.  
DR EMBL; L38592; AAA57556.1; JOINED.  
DR InterPro; IPR001046; Nramp.  
DR Pfam; PF01566; Nramp; 1.  
DR ProDom; PD001861; Nramp; 1.  
FT NON\_TER 1  
SQ SEQUENCE 72 AA; 7778 MW; 4B5EE8D1A67D8498 CRC64;

Query Match 2.6%; Score 7; DB 4; Length 72;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAARL 202  
Db 16 QRLAARL 22

RESULT 32  
Q075702 PRELIMINARY; PRT; 73 AA.  
AC Q075702;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)  
DE PROTEIN-TYROSINE-PHOSPHATASE, ISOFORM 3 (EC 3.1.3.48).  
GN ACPL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA Modesti A., Marzocchini R., Raugeri G., Chiti A., Sereni A.,  
RA Magherini F., Rampont G.;  
RT "Cloning, expression and characterisation of a new human low Mr  
RT phosphotyrosine protein phosphatase originating by alternative  
RT splicing";  
RL FEBS Lett. 0:0-0(0).  
DR EMBL; Y16846; CAA76416.1; -.  
DR HSSP; P24666; SPNT.  
DR InterPro; IPR000106; Low\_mwt\_PTPase.  
DR SMART; SM00226; LMWPC; 1.  
KW Hydrolase.  
SQ SEQUENCE 73 AA; 7660 MW; D261205427CBECF3 CRC64;

Query Match 2.6%; Score 7; DB 4; Length 73;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AVLEFLTG 15  
Db 54 AVLEFLTG 60

RESULT 33  
Q09DKM9 PRELIMINARY; PRT; 78 AA.  
AC Q09DKM9;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, last annotation update)  
DE HYPOTHEtical 9.0 KDA PROTEIN.  
OS Spodoptera frugiperda ascovirus 1.  
OC Viruses; dsDNA viruses, no RNA stage; Ascoviridae; Ascovirus.  
OX NCBI\_TaxID=113374;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20540044; PubMed=11086137;  
RA Stasiak K., Demattei M.V., Federici B.A., Bigot Y.;  
RT "Phylogenetic position of the Diadromus pulchellus ascovirus DNA  
RT polymerase among viruses with large double-stranded DNA genomes";  
RL J. Gen. Virol. 81:3059-3072(2000).  
DR EMBL; AJ279828; CAC19164.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 78 AA; 9012 MW; D3A91C118850EAE2 CRC64;

Query Match 2.6%; Score 7; DB 12; Length 78;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 SVTSTES 41

RESULT 34

O942Y3 PRELIMINARY; PRT; 83 AA.  
AC Q942Y3;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE P0492G09.14 PROTEIN.  
GN P0492G09.14.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
clone:P0492G09.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003266; BAB64192.1; -  
SQ SEQUENCE 83 AA; 8687 MW; B06A60DB1CD1814C CRC64;

## Query Match

Best Local Similarity 2.6%; Score 7; DB 10; Length 83;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 DRARAHV 180  
Db 19 DRARAHV 25

RESULT 35

O9DXP2 PRELIMINARY; PRT; 87 AA.  
AC Q9DXP2;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91CAS237750;  
RA Op de Coul E.L.M., Prins M., Cornelissen M., van der Schoot A.,  
Boufassa F., Brettle R., Hernandez-Aguado I., Schiffer V.,  
McMenamin J., Rezza G., Robertson R., Goudsmit J., Coutinho R.,  
Lukashov V.;  
RT "Using phylogenetic analysis to trace HIV-1 migration among western  
European IDUs seroconverting from 1984 to 1997.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF307254; AAG32422.1; -  
FT NON\_TER 1 87  
FT NON\_TER 1 87  
SQ SEQUENCE 87 AA; 9949 MW; 2103B8706FF2D927 CRC64;

## Query Match

Best Local Similarity 2.6%; Score 7; DB 15; Length 87;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 KLREQLG 89  
Db 77 KLREQLG 83

RESULT 36

O9EAM6 PRELIMINARY; PRT; 88 AA.  
AC O9EAM6;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE GP120 (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAB178R;  
RA Machuca R.A.R., Bogh M., Gerstoft J., Kvinesdal B., Pedersen C.,  
Obel N., Nielsen H., Nielsen C.;  
RT "HIV-1 subtypes in Denmark.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAB178R;  
RA Bogh M., Machuca R.A.R., Nielsen C.;  
RT "Subtype specific problems with Roche PCR.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ289516; CAC06500.1; -  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1 88  
FT NON\_TER 1 88  
SQ SEQUENCE 88 AA; 9966 MW; 9059197BD2C09D9A CRC64;

## Query Match

Best Local Similarity 2.6%; Score 7; DB 15; Length 88;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 KLREQLG 89  
Db 79 KLREQLG 85

RESULT 37

O9EAM4 PRELIMINARY; PRT; 88 AA.  
AC O9EAM4;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE GP120 (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PT527;  
RA Machuca R.A.R., Bogh M., Gerstoft J., Kvinesdal B., Pedersen C.,  
Obel N., Nielsen H., Nielsen C.;  
RT "HIV-1 subtypes in Denmark.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PT527;  
RA Bogh M., Machuca R.A.R., Nielsen C.;  
RT "Subtype specific problems with Roche PCR.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ289592; CAC06572.1; -  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.

KW AIDS: Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 88  
SQ SEQUENCE 88 AA; 9908 MW; 46C7E92C65B6BAAC CRC64;

Query Match 2.6%; Score 7; DB 15; Length 88;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KLRQLG 89  
DB 78 KLRQLG 84

RESULT 38

Q90RW4 PRELIMINARY; PRT; 88 AA.  
AC Q90RW4;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN 11  
RP SEQUENCE FROM N.A.  
RC STRAIN=CD\_206V3;  
RA Machado D.M., Delwart E.L., Diaz R.S., Oliveira C.F., Rawal B.D., Sullivan M., Gwin M., Clark K.A., Busch M.P.;  
RT "Use of the Dual Sensitive/Less-Sensitive (Detuned) EIA Strategy for Targeting Genetic Analysis of HIV-1 to Recently Infected U.S. Blood Donors."  
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF373756; AAK55377.1; -.  
DR EMBL; AF373756; AAK55377.1; -.  
FT NON\_TER 1  
FT NON\_TER 88  
SQ SEQUENCE 88 AA; 10037 MW; 575937C83EC5D42E CRC64;

Query Match 2.6%; Score 7; DB 15; Length 88;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KLRQLG 89  
DB 81 KLRQLG 87

RESULT 39

Q9BR82 PRELIMINARY; PRT; 89 AA.  
AC Q9BR82;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHETICAL 9.6 KDA PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN 11  
RP SEQUENCE FROM N.A.  
RC TISSUE=MUSCLE, RHABDOMYOSARCOMA;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC006438; AAH06438.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 89 AA; 9567 MW; 8908855FF350C3C2 CRC64;

Query Match 2.6%; Score 7; DB 4; Length 89;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLTAVL 11  
DB 26 VLTAVL 32

RESULT 40

Q9ID03 PRELIMINARY; PRT; 89 AA.  
AC Q9ID03;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ENVELOPE PROTEIN (FRAGMENT).  
GN ENV OR GP120.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN 11  
RP SEQUENCE FROM N.A.  
RC STRAIN=COL694;  
RX MEDLINE=99388930; PubMed=10461834;  
RA Navas M.C., Letourneur E., Gomas E., Boshell J., Saragosti S.;  
RT "Analysis of the V3 loop sequences from 12 HIV type-1 infected patients from Colombia, South America."  
RT AIDS Res. Hum. Retroviruses 15:1141-1144(1999).  
RL EMBL; Y10366; CAA71395.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120.1.  
KW AIDS: Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 89  
SQ SEQUENCE 89 AA; 9858 MW; 688C7D9A96FFE10F CRC64;

Query Match 2.6%; Score 7; DB 15; Length 89;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KLRQLG 89  
DB 70 KLRQLG 76

RESULT 41

Q78841 PRELIMINARY; PRT; 90 AA.  
AC Q78841;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ENVELOPE PROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN 11  
RP SEQUENCE FROM N.A.  
RC STRAIN=HOMOSEXUAL H36, SEROCONVERSION SAMPLE 1991;  
RA Kuiken C.L., Lukashov V., Baan E., Dekker J., Leunissen J.A.M., Goudsmit J.;  
RT "Evidence for limited intra-subject evolution of the V3 domain of the HIV-1 envelope in the Amsterdam population."  
RL AIDS 0:0-0(1996).  
DR EMBL; Z68059; CAA92075.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120.1.  
KW AIDS: Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 90





RA Wofls T.F., Zwart G., Bakker M., Goudsmit J.;  
RT "HIV-1 genomic RNA diversification following sexual and parenteral  
virus transmission."  
RL Virology 189:103-110(1992).  
DR EMBL; M91878; AAA73294.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 92 AA; 10351 MW; 99219A30601C373C CRC64;

Query Match 2.6%; Score 7; DB 15; Length 92;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KIREQLG 89  
|||||  
Db 77 KIREQLG 83

## RESULT 46

Q9WN08 PRELIMINARY; PRT; 94 AA.  
AC Q9WN08;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN GP120 (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M7;  
RX MEDLINE=20072106; PubMed=10606090;  
RA Zachar V., Zacharova V., Fink T., Thomas R.A., King B.R., Ebbesen P.,  
Jones T.B., Goustin A.S.;  
RT "Genetic analysis reveals ongoing HIV type 1 evolution in infected  
human placental trophoblast."  
RL AIDS Res. Hum. Retroviruses 15:1673-1683(1999).  
DR EMBL; AF150066; AAD43683.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 94 AA; 10731 MW; D92889598023B0DC CRC64;

Query Match 2.6%; Score 7; DB 15; Length 94;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KIREQLG 89  
|||||  
Db 79 KIREQLG 85

## RESULT 47

Q9IJU5 PRELIMINARY; PRT; 95 AA.  
AC Q9IJU5;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=1;  
RA Wang B., Wu H., Fang R.;  
RT "The nucleotide sequence analysis on the C2-V3 region of HIV1 strains  
in Shandong area."  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF220239; AAF74416.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 95 AA; 10557 MW; EADBF0B8C0BD7D2 CRC64;

Query Match 2.6%; Score 7; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KIREQLG 89  
|||||  
Db 74 KIREQLG 80

## RESULT 48

Q75278 PRELIMINARY; PRT; 95 AA.  
AC Q75278;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ENVELOPE GP120 V3 REGION (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95191036; PubMed=7884906;  
RA Sovia P., van Raast M., Gupta P., Balachandran R., Chao W., Itescu S.,  
McKinley G., Volsky D.J.;  
RT "Conservation of an intact human immunodeficiency virus type 1 vif  
gene in vitro and in vivo."  
RL J. Virol. 69:2557-2564(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Foley B.T.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U50621; AAA96040.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 95 AA; 10524 MW; 2114CB712DB449C2 CRC64;

Query Match 2.6%; Score 7; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KIREQLG 89  
|||||  
Db 77 KIREQLG 83

## RESULT 49

Q9WL10 PRELIMINARY; PRT; 96 AA.  
AC Q9WL10;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
RN [1]

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S14V3-3;  
RX MEDLINE=98445411; PubMed=9770526;  
RA Markham R.B., Wang W.C., Weistein A.E., Wang Z., Munoz A.,  
RA Templeton A., Margolick J., Vlahov D., Quinn T., Farzadegan H.,  
RA Yu X.F.;  
RT "Patterns of HIV-1 evolution in individuals with differing rates of  
RT CD4 T cell decline.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:12568-12573(1998).  
DR EMBL; AF089606; AAC70633.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1 1  
FT NON\_TER 96 96  
SQ SEQUENCE 96 AA; 10931 MW; C6F3888D0B4CAB28 CRC64;

Query Match 2.6%; Score 7; DB 15; Length 96;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KLREQLG 89  
| | | | | | | |  
Db 79 KLREQLG 85

RESULT 50  
Q9IDT9  
ID Q9IDT9 PRELIMINARY; PRT; 100 AA.  
AC Q9IDT9;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN GP120 (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20556821; PubMed=11103039;  
RA Machuca R.A.R., Schonning K., Hansen J., Fomsgaard A., Nielsen C.;  
RT "No association of HIV-1 envelope (C2-V3-C3) sequence pattern with  
RT long-term nonprogression.";  
RL J. Acquir. Immune Defic. Syndr. 25:103-108(2000).  
DR EMBL; AJ278207; CAB96355.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1 1  
FT NON\_TER 100 100  
SQ SEQUENCE 100 AA; 11347 MW; 29412D312341BD8C CRC64;

Query Match 2.6%; Score 7; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KLREQLG 89  
| | | | | | | |  
Db 85 KLREQLG 91

Search completed: September 22, 2002, 12:26:53  
Job time: 238 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:06:35 ; Search time 29.35 Seconds  
(without alignments)  
1010.450 Million cell updates/sec

Title: US-09-803-918A-2  
Perfect score: 267  
Sequence: 1 MKAAVLTIAVLFLTGSQARH.....SFKVSFLSALEEYTKKLTNQ 267

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : A\_Geneseq\_032802:\*

1:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	267	100.0	267	7 AAP61079	Assumed human apol
2	267	100.0	267	9 AAP82128	Entire human prepr
3	267	100.0	267	14 AAR34032	Sequence of apo AI
4	267	100.0	267	16 AAR72705	Human apo A-I incl
5	267	100.0	267	20 AAY18675	Human apolipoprote
6	267	100.0	267	22 AAB47620	Full length Apo-AI
7	249	93.3	268	9 AAP80668	Recombinant human
8	243	91.0	243	9 AAP81082	Sequence of mature
9	193	72.3	244	22 AAU28184	Novel human secret
10	191	71.5	299	22 AAU33170	Novel human secret
11	174	65.2	267	18 AAW08602	Human apolipoprote

12	172	64.4	264	15	AAR56863	Apo-lipoprotein AI
13	172	64.4	264	15	AAR56864	Apo-lipoprotein AI
14	143	53.6	221	22	AAU29835	Novel human secret
15	97	36.3	166	22	AAU28372	Novel human secret
16	92	34.5	119	22	AAU30468	Novel human secret
17	68	25.5	120	22	AAU30267	Novel human secret
18	68	25.5	120	22	AAU30469	Novel human secret
19	48	18.0	318	22	AAU30268	Novel human secret
20	48	18.0	359	22	AAU30470	Novel human secret
21	42	15.7	42	13	AAR20164	Apo AI polypeptide
22	26	9.7	26	22	AAG62609	Apolipoprotein fra
23	26	9.7	134	22	AAU29834	Novel human secret
24	25	9.4	151	22	AAO02278	Human polypeptide
25	22	8.2	22	10	AAP92072	Apolipoprotein AI
26	22	8.2	22	22	AAG62608	Apolipoprotein fra
27	22	8.2	154	22	AAO12095	Human polypeptide
28	21	7.9	21	10	AAP90956	Apo AI epitope. S
29	21	7.9	32	13	AAP92073	Apolipoprotein AI
30	19	7.1	19	10	AAP92073	Apolipoprotein AI
31	18	6.7	18	20	AAV18676	Peptide SEQ ID NO:
32	18	6.7	38	22	AAO08726	Human polypeptide
33	17	6.4	17	17	AAR98474	Anti-arteriosclero
34	16	6.0	16	14	AAR34033	Sequence of apo AI
35	16	6.0	16	16	AAR72706	Human apo A-I frag
36	16	6.0	16	20	AAV27065	Peptide Seq ID NO:
37	16	6.0	16	22	AAG62607	Apolipoprotein fra
38	15	5.6	16	7	AAP60986	Sequence of immuno
39	13	4.9	14	7	AAP60985	Sequence of immuno
40	12	4.5	12	10	AAP92076	Apolipoprotein AI
41	12	4.5	12	20	AAV41961	Rheumatoid arthritis
42	12	4.5	12	20	AAV42027	Rheumatoid arthritis
43	12	4.5	12	22	AAAB87196	Breast-cancer asso
44	12	4.5	13	7	AAP60987	Sequence of immuno
45	12	4.5	22	22	AAAB87236	Breast-cancer asso
46	11	4.1	11	20	AAV41957	Rheumatoid arthritis
47	11	4.1	11	20	AAV42023	Rheumatoid arthritis
48	11	4.1	11	22	ABBS6060	Vascular dementia-
49	11	4.1	11	22	AAU24926	Schizophrenia-Asso
50	11	4.1	11	22	AAU15270	Schizophrenia-Asso
51	11	4.1	11	22	AAAB87194	Breast-cancer asso
52	11	4.1	11	22	AAAB87195	Breast-cancer asso
53	11	4.1	11	22	AAAB87222	Breast-cancer asso
54	11	4.1	11	22	AAAB87234	Breast-cancer asso
55	11	4.1	11	22	AAAB87235	Breast-cancer asso
56	11	4.1	21	22	AAAB87221	Breast-cancer asso
57	10	3.7	10	20	AAV41956	Rheumatoid arthritis
58	10	3.7	10	20	AAV42022	Rheumatoid arthritis
59	10	3.7	10	22	AAAB87197	Breast-cancer asso
60	10	3.7	10	22	AAAB87223	Breast-cancer asso
61	9	3.4	9	20	AAV41955	Rheumatoid arthritis
62	9	3.4	9	20	AAV41960	Rheumatoid arthritis
63	9	3.4	9	20	AAV42021	Rheumatoid arthritis
64	9	3.4	9	20	AAV42026	Rheumatoid arthritis
65	9	3.4	9	22	AAAB87193	Breast-cancer asso
66	9	3.4	9	22	AAAB87220	Breast-cancer asso
67	9	3.4	9	22	AAAB87233	Breast-cancer asso
68	9	3.4	13	15	AAR57883	ICAT minimum bindi
69	9	3.4	15	10	AAP92074	Apolipoprotein AI
70	9	3.4	16	10	AAP92075	Apolipoprotein AI
71	8	3.0	8	18	AAW32556	Amyloidogenic sequ
72	8	3.0	8	20	AAV41959	Rheumatoid arthritis
73	8	3.0	8	20	AAV42025	Rheumatoid arthritis
74	8	3.0	15	11	AAR07703	Fusion protein com
75	8	3.0	317	19	AAW37932	Amino acid sequenc

#### ALIGNMENTS

RESULT 1  
ID AAP61079  
XX AAP61079 standard; Protein; 267 AA.



AC AAP61079;  
XX  
DT 07-OCT-1991 (first entry)  
XX  
DE Assumed human apolipoprotein A-1 derivative gene product.  
XX  
KW Hyperlipaemia; arteriosclerosis.  
XX  
OS Homo sapiens.  
XX  
PN JP61096998-A.  
XX  
PD 15-MAY-1986.  
XX  
PF 16-OCT-1984; 84JP-0216988.  
XX  
PR 16-OCT-1984; 84JP-0216988.  
XX  
PA (MITU ) MITSUBISHI CHEM IND KK.  
XX  
DR WPI; 1986-165025/26.  
DR N-PSDB; AAN60886.  
XX  
PT Human apolipoprotein A-1 (deriv.) prepn. - by providing DNA  
PT fragment in cloning site downstream of expression vector promoter  
PT and introducing into host microorganism.  
XX  
PS Disclosure; Fig 2; 9pp; Japanese.  
XX  
CC The human apolipoprotein may be produced by a suitable transformed  
CC host, it is effective in treating hyperlipaemia and arteriosclerosis.  
XX  
SQ Sequence 267 AA;

Query Match 100.0%; Score 267; DB 7; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.2e-241;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKAAVLTAVLELTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFECS 60  
DB 1 mkaavltavlflltgsqarhfwqdeppspwdrvkdlatvyvdvlkdsgrdyvsqfegs 60  
QY 61 ALGKQNLKLLDNWDSVSTFSKLRQGLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnwdsvstfsklreqlpgvtqefwdnleketeglrqemskdleevkak 120  
QY 121 VQPYLDFFQKKQWQEMELYRQKVEPLRAELQEGARQKLTHELQKLSPLGEMRDRARAHV 180  
DB 121 vqpyldffqkkwqeemelyrqkveplraelqegarqklhelqeklsplgeemrdrarahv 180  
QY 181 DALRTHLAPYSDELQRRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapysdelqrarlaarlealkengarlarlaeyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESFQVSLSALEEYTKKLTNQ 267  
DB 241 glpvllesfksvlsaleeytkkltntq 267

RESULT 2  
AAP82128  
ID AAP82128 standard; protein; 267 AA.  
XX  
AC AAP82128;  
XX  
DT 24-OCT-1990 (first entry)  
XX  
DE Entire human preproapoprotein A1.  
XX  
KW human preproapoprotein A1; high density lipoprotein deficiency; ss.  
XX  
OS synthetic.  
1

XX Key Location/Qualifiers  
FH Peptide 1..18  
FT Peptide /label=precursor  
FT Peptide 19..24  
FT Protein /label=propeptide  
FT Protein 25..267  
FT Protein /label=mature apoprotein  
XX  
PN EP293357-A.  
XX  
PD 30-NOV-1988.  
XX  
PF 24-MAY-1988; 88EP-0870095.  
XX  
PR 28-MAY-1987; 87GB-0012540.  
XX  
PA (UNIO ) UCB SA.  
XX  
PI Bollen A, Gobert J, Wulfert E;  
XX  
DR WPI; 1988-339891/48.  
DR N-PSDB; AAN82064.  
XX  
PT New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures  
XX  
PS Disclosure; ; p; French.  
XX  
CC The cDNA 878bp fragment encoding preproapoprotein A1 was detected  
CC in clone pUB1609 derived from human liver cells.  
XX  
SQ See also AAN81258.  
SQ Sequence 267 AA;

Query Match 100.0%; Score 267; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.2e-241;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKAAVLTAVLELTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFECS 60  
DB 1 mkaavltavlflltgsqarhfwqdeppspwdrvkdlatvyvdvlkdsgrdyvsqfegs 60  
QY 61 ALGKQNLKLLDNWDSVSTFSKLRQGLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnwdsvstfsklreqlpgvtqefwdnleketeglrqemskdleevkak 120  
QY 121 VQPYLDFFQKKQWQEMELYRQKVEPLRAELQEGARQKLTHELQKLSPLGEMRDRARAHV 180  
DB 121 vqpyldffqkkwqeemelyrqkveplraelqegarqklhelqeklsplgeemrdrarahv 180  
QY 181 DALRTHLAPYSDELQRRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapysdelqrarlaarlealkengarlarlaeyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESFQVSLSALEEYTKKLTNQ 267  
DB 241 glpvllesfksvlsaleeytkkltntq 267

RESULT 3  
AAR34032  
ID AAR34032 standard; Protein; 267 AA.  
XX  
AC AAR34032;  
XX  
DT 13-AUG-1993 (first entry)  
XX  
DE Sequence of apo A1.  
XX  
KW Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.  
XX

OS Homo sapiens.  
XX  
PN WO9307165-A.  
XX  
PD 15-APR-1993.  
XX  
PF 09-OCT-1992; 92WO-US08634.  
XX  
PR 09-OCT-1991; 91US-0774633.  
PR 08-OCT-1992; 92US-055555.  
PR 28-JUN-1992; 92US-0901706.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;  
XX  
DR WPI: 1993-134378/16.  
DR N-PSDB; AAQ40030.  
XX  
PT Polypeptide mimic of native apo B-100 and native apo A-I - useful  
PT in assays for LDL and HDL in plasma samples  
XX  
PS Claim 19; Pages 105-106; 137pp; English.  
XX  
CC The inventors claim a portion of the polypeptide contg. apo B-100  
CC that immunoreacts with antibodies secreted by the hybridoma MB47  
CC having ATCC Accession No. 8746. Polypeptides specifically claimed  
CC include residues 217-297, 216-310, 216-331, 216-352, 216-377, 1-377,  
CC 205-297, 173-297, 140-297. DNA sequences encoding the polypeptides  
CC are also claimed. Also claimed are a fusion polypeptide that  
CC contains: (a) a first amino acid residue sequence up to 250 residues  
CC in length that includes residues 120-135 of apo A-I, (b) a second  
CC amino acid residue sequence up to 375 residues in length that  
CC includes residues 217-297 of apo B-100 and DNA encoding it.  
XX  
SQ Sequence 267 AA;  
  
Query Match 100.0%; Score 267; DB 14; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.2e-241;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MKA AVLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGS 60  
DB 1 mkaavltlavlfiltsgarhfwqdeppqspwdrvkdlatvyvdvldksgrdyvsqfegs 60  
  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLPVTOEFWMDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnwdsvststfsklreqlpgvtqefwmdnleketeglrqemskdleevkak 120  
  
QY 121 VOPYLDDFOKKWQEMELLYRQKVEPLRAELQEGAROKLHELQEKISPLGEEMDRARAHV 180  
DB 121 vqpylddfqkkwqeemellyrqkveplraelqegaroklhelqekisplgeemdrarahv 180  
  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapysdelrqlraarlealkengarlarlaeyhakatehstlsekakpaledlrq 240  
  
QY 241 GLLPVLESFKVVSFLSALBEYTKKLTNQ 267  
DB 241 gllpvlesfkvsflsaleeytkkltntq 267  
  
RESULT 4  
ID AAR72705 standard; Protein; 267 AA.  
XX  
AC AAR72705;  
XX  
DT 31-OCT-1995 (first entry)  
XX  
DE Human apo A-I including signal and propeptide sequences.  
XX

KW Apo A-I; LDL cholesterol; low density lipoprotein; lipid.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FH Peptide 1..18  
FT /label= presignal  
FT Peptide 19..24  
FT /label= propeptide  
FT Peptide 120..135  
FT /label= claimed  
FT /note= "as part of fusion polypeptide"  
FT Peptide 19..240  
FT /label= claimed  
FT /note= "as part of fusion polypeptide"  
XX  
PN US5408038-A.  
XX  
PD 18-APR-1995.  
XX  
PF 09-OCT-1991; 91US-0774633.  
XX  
PR 09-OCT-1991; 91US-0774633.  
PR 18-JUN-1992; 92US-0901706.  
PR 08-OCT-1992; 92US-0959946.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;  
XX  
DR WPI: 1995-161146/21.  
DR N-PSDB; AAQ89634.  
XX  
PT New apo:lipoprotein B-100 peptide(s) and fusion peptide(s) - used  
PT in assay systems for detecting LDL and HDL cholesterol levels in  
PT body fluids.  
XX  
PS Claim 10; Fig 2; 41pp; English.  
XX  
CC AAA89634 and AAR72705 depict the AA sequence of human apo A-I and its  
CC corresp. cDNA, including presignal residues and propeptide  
CC residues, according to Seilhamer et al., DNA 3(4):309 (1984).  
CC A dispersible apo A-I/B-100 fusion polypeptide is claimed which  
CC contains a first AA sequence of apo A-I and that includes at  
CC least AA sequence positions 120-135 (see AAR72606) and which reacts  
CC with pan anti-apo A-I antibodies such as: AI-4 ATCC HB8744; AI-7  
CC ATCC HB 8745; AI-9 ATCC HB 8741; AI-10 ATCC HB 9200; AI-11 ATCC  
CC HB 9201; AI-12 ATCC HB 9202; AI-13 ATCC HB 9203; AI-14 ATCC HB  
CC 9204; AI-18 ATCC HB 9507.  
XX  
SQ Sequence 267 AA;  
  
Query Match 100.0%; Score 267; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.2e-241;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MKA AVLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGS 60  
DB 1 mkaavltlavlfiltsgarhfwqdeppqspwdrvkdlatvyvdvldksgrdyvsqfegs 60  
  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLPVTOEFWMDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnwdsvststfsklreqlpgvtqefwmdnleketeglrqemskdleevkak 120  
  
QY 121 VOPYLDDFOKKWQEMELLYRQKVEPLRAELQEGAROKLHELQEKISPLGEEMDRARAHV 180  
DB 121 vqpylddfqkkwqeemellyrqkveplraelqegaroklhelqekisplgeemdrarahv 180  
  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapysdelrqlraarlealkengarlarlaeyhakatehstlsekakpaledlrq 240



FT /label= AFTI  
FT /note= "13 kD N-terminal fragment"  
FT 156..267  
FT /label= AFTI  
FT /note= "13 kD C-terminal fragment"

WO200168852-A2.

20-SEP-2001.

13-MAR-2001; 2001WO-US07826.

13-MAR-2000; 2000US-189008P.

(AMGE-) AMGEN INC.

Edwards CK, Burger D, Dayer J, Kohno T;

WPI; 2001-596908/67.

N-PSDB; AAH43623.

Apo-A-I fragment T-cell activation inhibitor (AFTI) polynucleotides,  
useful for treating, diagnosing, ameliorating diseases associated with  
IL-1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease  
and asthma -

Claim 1; Fig 1A; 132pp; English.

This sequence shows full length apolipoprotein (Apo-AI). Fragments  
of Apo-AI may be used as Apo-A-I fragment T-cell activation inhibitors  
(AFTI). These fragments are selected from an 18 kD N-terminal fragment  
(amino acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144)  
and a 13 kD C-terminal fragment (amino acids 156-267). The AFTI  
polypeptides and polynucleotides are useful for regulating T-cell  
mediated activation of monocytes and for treating, diagnosing,  
ameliorating diseases associated with IL-1 and/or TNF activity.  
The diseases are acute pancreatitis, Alzheimer's disease, asthma,  
cancer, fever, inflammatory bowel disease, ischemia, multiple  
sclerosis, osteoporosis, Parkinson's disease or psoriasis. Numerous  
examples of other diseases are given in the specification.  
The AFTI nucleic acids are useful as hybridization probes in diagnostic  
assays to test for the presence of an AFTI or Apo-A-I DNA in mammalian  
tissue or bodily fluid samples.

Sequence 267 AA;

Query Match 100.0%; Score 267; DB 22; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.2e-241;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRAVLTAVLFLTGSQARHFWQDEPPSPWDVRYKDLATVYVDVLDKSGRDYVSQFEGS 60  
Db 1 mkaavltlavflitgsqarhfwqdeppspwdrvkdlatvyvdvldksgrdyvsqfegs 60  
QY 61 ALGKQLNLKLLDNMDSVTSTFSKLRQGLPVTQEFWNLKETEGLRQEMSKDLEEVKAK 120  
Db 61 alqkqlnlkldnmdsvtstfslkrlqglpvtqefwlnlketeglrqemskdleevkak 120  
QY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHLEQKLSPIGEMRDRARAHV 180  
Db 121 vqpylddfqkkwqemellyrqkveplraelqegarqklhelqeklspligemrdrarahv 180  
QY 181 DALRTHLAPYSDELQRLARLEALKENGARGLAELYHAKATEHLSTLSEKAKPALEDLRQ 240  
Db 181 dalrthlapysdelrqlaarlealkengarglaelyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESFVKVSTLSALEEYTKKINTQ 267  
Db 241 glipvlesfksvstlsaleeytkkintq 267

RESULT 7

AAP80668  
ID AAP80668 standard; protein; 268 AA.  
XX  
XX AAP80668;  
AC  
XX 24-OCT-1990 (first entry)  
DT  
XX Recombinant human preproapoprotein A1.  
DE  
XX human preproapoprotein A1; high density lipoprotein deficiency; ss.  
KW  
XX synthetic.  
OS  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT /label=precursor  
FT Peptide 19..25  
FT /label=proprotein  
FT Protein 26..268  
FT /label=mature apoprotein A1

EP293357-A.

30-NOV-1988.

24-MAY-1988; 88EP-0870095.

28-MAY-1987; 87GB-0012540.

(UNIO ) UCB SA.

Bollen A, Gobert J, Wulfert E;

WPI; 1988-339891/48.

N-PSDB; AAN81258.

New DNA encoding human preproapoprotein A1 -  
modified to eliminate hairpin structures

Claim 1; Page 12; 25pp; French.

Met at posn 19 is inserted as an extra amino acid c.f. wild-type  
protein.  
The DNA fragment used to replace the wild-type sequence encoding  
amino acids -6 to +14 (= Arg 20 to Leu 39) modifies codons  
corresponding to wild-type amino acids -6, -1, 1, 3, 4, 5, 6, 7,  
10, 11 and 14. The changed codons still encode the same amino acids  
as in the wild-type protein but reduce formation of secondary  
structures in mRNA.  
See also AAN82064.

Sequence 268 AA;

Query Match 93.3%; Score 249; DB 9; Length 268;  
Best Local Similarity 100.0%; Pred. No. 8.4e-225;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 RHFWQDEPPSPQSPWDRVVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLLDNMDSVT 78  
Db 20 rhfwqdeppspqspwdrvkdlatvyvdvldksgrdyvsqfegsalgkqlnlkldnmdsvt 78  
QY 79 STFSKLRQGLPVTQEFWNLKETEGLRQEMSKDLEEVKAKVOPYLDDFQKKWQEMEL 138  
Db 80 stfslkrlqglpvtqefwlnlketeglrqemskdleevkakvopylddfqkkwqemel 138  
QY 139 YRQKVEPLRAELQEGARQKLHLEQKLSPIGEMRDRARAHVDALRTHLAPYSDELQRL 198  
Db 140 yrqkveplraelqegarqklhelqeklspligemrdrarahvdalrthlapysdelrql 198  
QY 199 AARLEALKENGARGLAELYHAKATEHLSTLSEKAKPALEDLRQGLPVLSEFVKVSTLSALE 258  
Db 200 aarlealkengarglaelyhakatehlstlsekakpaledlrqglpvlsefksvstlsale 258



QY 259 EYTKKLTQ 267  
1111111111  
Db 260 eytkkltq 268

## RESULT 8

AAP81082  
ID AAP81082 standard; protein; 243 AA.

XX  
AC AAP81082;

DT 14-JAN-1991 (first entry)

XX Sequence of mature human apolipoprotein AI (apoAI).

XX Atherosclerosis; therapy; cardiovascular disease.

XX Homo sapiens.

XX WO8803166-A.

XX 05-MAY-1988.

XX 21-OCT-1987; 87WO-EP00621.

XX 23-OCT-1986; 86GB-0025435.

XX (FARM ) FARMITALIA C ERBA SPA.

XX Lorenzetti R, Monaco L, Soria M, Palomba R, Isacchi A, Sarmientos P;

XX MPI; 1988-133240/19.

XX N-PSDB; AAN80243.

XX Recombinant human apo:lipoprotein AI -  
PT used to lower plasma cholesterol and/or tri glyceride levels and  
PT to combat atherosclerosis and cardiovascular diseases

XX Disclosure; Fig 1; 51pp; English.

XX The protein comprising apo AI genetic variants may be used to lower  
CC plasma cholesterol and/or triglyceride levels. They may also be used to  
CC combat atherosclerosis and cardiovascular diseases such as coronary  
CC heart disease. Prefd. proteins are Met-apo AI, Met-apo AI-T6, Met-apo  
CC AI-MI and Met-apo AI-T6/MI.

XX Sequence 243 AA;

Query Match 91.0%; Score 243; DB 9; Length 243;

Best Local Similarity 100.0%; Pred. No. 3.2e-219;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DEPPQSPWDRAVKDLATVYVDYVKDSGRDYSQFEGSALGKQLNLKLDNMWDSVTSTFSL 84

Db 1 deppqspwdrvkdlatvdyvdkdsgdyvsqfegsalgkqlnlkldnwdsvtstfsl 60

QY 85 REQLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 144

Db 61 reqlgptvqefwdnleketeglrqemskdleevkavqpylddfqkxwqeemelyrqkve 120

QY 145 PLRAELQEGAROKLHELQEKLSPIGEMDRARAHVVALRTHLAPYSDELQRLARLEA 204

Db 121 plraelqegarqklhelqeklspligemdrarahvvalrthlapysdelqrlailea 180

QY 205 LKENGARLAEXHAKATEHLSTSEKAKPALEDLRQGLLPVLESFVSLSEYTKKL 264

Db 181 lkengarlaexhakatehlstlsekakpaledlrqgllpvlsefksflsaeyletkkl 240

QY 265 NTQ 267

Db 241 ntq 243

## RESULT 9

AAU28184  
ID AAU28184 standard; Protein; 244 AA.

XX  
AC AAU28184;

DT 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 353.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US04942.

XX 07-MAR-2000; 2000US-0519705.

XX 19-MAY-2000; 2000US-0574454.

XX 17-JUN-2000; 2000US-0596193.

XX 14-JUL-2000; 2000US-0616847.

XX 19-SEP-2000; 2000US-0665363.

XX 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

XX Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX MPI; 2001-589934/66.

XX N-PSDB; AAS45084.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries  
PT prepared from various human tissues, for diagnosis and treatment of  
PT cancer, neurological, inflammatory, and autoimmune disorders -

XX Example 5; SEQ ID No 353; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)  
CC and polynucleotides (II). (I) and (II) are useful for treating  
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
CC involved in increasing haematopoiesis, stem cell survival, bone growth  
CC and remodeling. (I), (II) and modulators of (II) are useful for  
CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
CC creating transgenic animals useful for studying the in vivo activities of  
CC the polypeptide as well as for studying modulators of the polypeptides.  
CC (I) induces the proliferation of neural cells and regeneration of nerve  
CC and brain tissue and is useful for the treatment of central and  
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
CC activity, regulation of haematopoiesis and is useful for treating myeloid  
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
CC tissue growth, and in tissue repair, healing of burns, incisions,  
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative  
CC disorders, or periodontal disease. Furthermore, (I) is also useful for  
CC gut protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues, various immune deficiencies and



disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention.

Sequence 244 AA;

Query Match 72.3%; Score 193; DB 22; Length 244;  
Best Local Similarity 100.0%; Pred. No. 2e-172;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 DSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQE 134  
Db 52 dsvstfskrlreqlgpvtqefwdnleketeglrqemskdleevkavqpylddfqkkwqe 111  
QY 135 EMELRYQKVEPLRAELQEGARQKLHLEQKLSPLGEMDRARAHVDALRTHLAPYSDEL 194  
Db 112 emelyrqkveplraelqegarqklhelqeklsplgeemdrarahvdalrthlapydel 171  
QY 195 RQRLAARLEALKENGARLAETHAKATEHLSTLSEKAPALEDLRQGLLPVLESFQVSVFL 254  
Db 172 rqrtaarleaalkengarlaeyhakatehlstlsekapaledlrqglpvlsefsvsfl 231  
QY 255 SALEEYTKKLNTQ 267  
Db 232 saleeytkklnltq 244

RESULT 10

AAU33170  
ID AAU33170 standard; Protein; 299 AA.

AC AAU33170;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3661.

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

WPI; 2001-611725/70.

Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy -

Claim 20; Page 718; 765pp; English.

The invention relates to novel human secreted polypeptides. The

polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid sequences of novel human secreted proteins of the invention.

Sequence 299 AA;

Query Match 71.5%; Score 191; DB 22; Length 299;  
Best Local Similarity 100.0%; Pred. No. 1.8e-170;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSAARHFQWQDEPPQSPMDRVKDLATVYVDYLDKSGRDYVSQFEES 60  
Db 33 mkaavltavlitgsgarhfwqdeppqspmdrvkdlatvyvdylkdsgrdyvsqfegs 92  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
Db 93 algkqlnlkldnwdsvststfskrlreqlgpvtqefwdnleketeglrqemskdleevkak 152  
QY 121 VQPYLDDFQKKWQEMELRYRQKVEPLRAELQEGARQKLHLEQKLSPLGEMDRARAHV 180  
Db 153 vqpylddfqkkwqeemelyrqkveplraelqegarqklhelqeklsplgeemdrarahv 212  
QY 181 DALRTHLAPYS 191  
Db 213 dalrthlapyys 223

RESULT 11

AAW08602  
ID AAW08602 standard; Protein; 267 AA.

AC AAW08602;

DT 04-SEP-1997 (first entry)

DE Human apolipoprotein A-1 variant "Paris" protein sequence.

KW Human; apolipoprotein A-1; variant; mutation; serum lipid; cholesterol;

KW high density lipoprotein; triglyceride; symptom; cardiovascular disease;

KW atherosclerosis; restenosis; myocardial infarct; angina; death; stroke;

KW cardiac decompensation; metabolic deficit.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 1..24 /note= "prepro peptide"

FT Region 172..178 /note= "this sequence which contains the mutated amino acid residue is claimed - claim 2"

FT Misc-difference 175 /note= "changed from Arg residue in wild type protein, due to a C to T transition mutation"

PN WO9637608-A1.

PD 28-NOV-1996.

PF 20-MAY-1996; 96WO-FR00747.

XX 22-MAY-1995; 95FR-0006061.  
PR  
XX  
PA (INSP ) INST PASTEUR LILLE.  
PA (RHON ) RHONE-POULENC RORER SA.  
PA (UYPA-) UNIV CURIE PARIS VI P & M.  
XX  
PI Benoit P, Bruckert E, Deneffe P, Duverger N, Fruchart J;  
PI Luc G, Turping, Assmann G, Funke H;  
XX  
DR WPI; 1997-021218/02.  
DR N-PSDB; AAT43691.  
XX  
PT New variant of human apoA-1 with Cys at position 151 - has  
PT anti-atherogenic activity for treatment and prevention of  
PT cardiovascular disease  
XX  
PS Claim 2; Page -: 58pp; French.  
XX  
CC This is the amino acid sequence of a human apolipoprotein A-1 variant  
CC designated the "Paris" variant which has a Cys replacing the Arg residue  
CC at position 151. The substitution is generated by a mutation of  
CC the C nucleotide at position 523 in the wild type gene to a T residue,  
CC changing the encoded residue from an Arg to a Cys. The gene was isolated  
CC from a patient with an unusual pattern of serum lipids i.e. low levels of  
CC apoA-1 and high density lipoprotein (HDL)-cholesterol and high  
CC triglycerides, but showing symptoms of atherosclerosis. The new variant  
CC protein has anti-atherogenic activity so is useful for treatment and  
CC prevention of cardiovascular diseases such as atherosclerosis,  
CC restenosis, myocardial infarct, angina, sudden death, stroke and cardiac  
CC decompensation, or more generally any condition involving genetic or  
CC metabolic deficit of apoA-1.  
CC Note: this sequence is not given in the specification but is generated  
CC from the wild type apoA-1 gene disclosed in the specification and has  
CC the appropriate amino acid changed.  
XX  
SQ Sequence 267 AA;  
QY  
Db  
QY 1 MRAAVLTAVLFLTGSQARHFWQODEPPQSPWDVRVKDLATVYVDVLKDSGRDYVSQFEGS 60  
Db 1 mkaavltavflftgsqarhfwqodeppspwdrvkdlatvyvdvlkdsgrdyvsqfegs 60  
QY 61 ALGKQLNLKLLDNWDSVTSFESKLRQLGSPVTOEFWNLLEKTEGLRQEMSKDLEEVKAK 120  
Db 61 algkqlnlklldnwdsvtstfeskrlrglspvtgefwndlleketeglrqemskdleevkak 120  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMRD 174  
Db 121 vqpylddfqkkwqeemelyrqkveplraelqegarqklhelqeklsplgeemrd 174  
RESULT 12  
AAR56863  
ID AAR56863 standard; Protein; 264 AA.  
XX  
AC AAR56863;  
XX  
DT 26-JAN-1995 (first entry)  
XX  
DE Apo-lipoprotein AI-M.  
XX  
KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid pKP683;  
KW vector; atherosclerosis; cardiovascular disease.  
XX  
OS Homo sapiens.  
XX  
PN WO9413819-A.  
XX

PD 23-JUN-1994.  
XX  
XX  
PF 09-DEC-1993; 93WO-SE01061.  
XX  
PR 11-DEC-1992; 92SE-0003753.  
XX  
PA (KABI ) KABI PHARMACIA AB.  
XX  
XX  
PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
PI Sejlitz T;  
XX  
DR WPI; 1994-217892/26.  
DR N-PSDB; AAQ68357.  
XX  
PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.  
XX  
PS Disclosure; Fig. 3; 33pp; English.  
XX  
CC Plasmid pKP683 encodes human apo-lipoprotein AI-M in E. coli. The  
CC NotI-HindIII segment of pKP683 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.  
XX  
SQ Sequence 264 AA;  
QY  
Db  
QY 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSFESKL 84  
Db 22 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlklldnwdsvtstfsl 81  
QY 85 REQLGPVTQEFWNLLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVE 144  
Db 82 reqlgpvtqefwndlleketeglrqemskdleevkakvpylddfqkkwqeemelyrqkve 141  
QY 145 PLRAELOEGAROKLHELQEKLSPLGEMRDRARAHVDAIRTHLAPYSDELRLQ 196  
Db 142 plraelqegarqklhelqeklsplgeemrdrarahvdairthlapydsdelrq 193  
RESULT 13  
AAR56864  
ID AAR56864 standard; Protein; 264 AA.  
XX  
AC AAR56864;  
XX  
DT 26-JAN-1995 (first entry)  
XX  
DE Apo-lipoprotein AI-M.  
XX  
KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid pKP764;  
KW vector; atherosclerosis; cardiovascular disease.  
XX  
OS Homo sapiens.  
XX  
PN WO9413819-A.  
XX  
PD 23-JUN-1994.  
XX  
PF 09-DEC-1993; 93WO-SE01061.  
XX  
PR 11-DEC-1992; 92SE-0003753.  
XX  
PA (KABI ) KABI PHARMACIA AB.  
XX  
PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
PI Sejlitz T;

XX WPI; 1994-217892/26.  
DR N-PSDB; AA068358.  
XX  
PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.  
XX  
PS Disclosure; Fig. 4; 33pp; English.  
XX  
CC Plasmid PKP764 encodes human apo-lipoprotein AI-M in E. coli. The  
CC Moti-HindIII segment of PKP764 and the deduced amino acid sequence  
CC of the translated Apo AI-M protein were determined. The plasmid  
CC provides high yields of extracellular Apo AI-M.  
XX  
SQ Sequence 264 AA;  
  
Query Match 64.4%; Score 172; DB 15; Length 264;  
Best Local Similarity 100.0%; Pred. No. 9.9e-153;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 25 DEPPQSPWDRVKDLATVYVDVLKDSGRVYSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
DB 22 depqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalsgkqlnlkldnwdsvststfsl 81  
QY 85 REQLGPVTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEMELYRQKVE 144  
DB 82 reqlgpvtqefwdnleketegrlqemskdleevkavqpylddfqkkwgeemelyrqkve 141  
QY 145 PLRAELQEGARQKLHLEQKLSPLGEMRDRAHVDALRTHLAPYSDELRLQ 196  
DB 142 plraelqegarqklhelqeklsplgeemrdarahvdalrthlapysdelrlq 193  
  
RESULT 14  
AAU29835  
ID AAU29835 standard; Protein; 221 AA.  
XX  
AC AAU29835;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #326.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 199; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 221 AA;  
  
Query Match 53.6%; Score 143; DB 22; Length 221;  
Best Local Similarity 100.0%; Pred. No. 1.2e-125;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 88 IGPVTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEMELYRQKVEPLR 147  
DB 79 lgpvtqefwdnleketegrlqemskdleevkavqpylddfqkkwgeemelyrqkveplr 138  
QY 148 AELQEGARQKLHLEQKLSPLGEMRDRAHVDALRTHLAPYSDELRLQRLARLEALKE 207  
DB 139 aelqegarqklhelqeklsplgeemrdarahvdalrthlapysdelrlqrlaarlalke 198  
QY 208 NGGARLAELYHAKATEHSTLSEK 230  
DB 199 nggarlaelyhakatehstlsek 221  
  
RESULT 15  
AAU28372  
ID AAU28372 standard; Protein; 166 AA.  
XX  
AC AAU28372;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secretory protein, Seq ID No 729.  
XX  
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
KW fertility; analgesic; pain; antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO200166689-A2.  
XX  
PD 13-SEP-2001.  
XX  
PF 05-MAR-2001; 2001WO-US04942.  
XX  
PR 07-MAR-2000; 2000US-0519705.  
PR 19-MAY-2000; 2000US-0574454.  
PR 17-JUN-2000; 2000US-0596193.  
PR 14-JUL-2000; 2000US-0616847.  
PR 19-SEP-2000; 2000US-0665363.  
PR 20-OCT-2000; 2000US-0693267.  
XX  
PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;  
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;  
XX WPI; 2001-589934/66.  
DR N-PSDB; AAS45272.  
XX  
PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
PT prepared from various human tissues, for diagnosis and treatment of  
PT cancer, neurological, inflammatory, and autoimmune disorders -  
XX  
PS Example 2; SEQ ID No 729; 107pp; English.  
XX  
CC The invention relates to novel isolated human secreted polypeptides (I)  
CC and polynucleotides (II). (I) and (II) are useful for treating  
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
CC involved in increasing haematopoiesis, stem cell survival, bone growth  
CC and remodeling. (I), (II) and modulators of (II) are useful for  
CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
CC creating transgenic animals useful for studying the in vivo activities of  
CC the polypeptide as well as for studying modulators of the polypeptides.  
CC (I) induces the proliferation of neural cells and regeneration of nerve  
CC and brain tissue and is useful for the treatment of central and  
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
CC activity, regulation of haematopoiesis and is useful for treating myeloid  
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
CC tissue growth, and in tissue repair, healing of burns, incisions,  
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative  
CC disorders, or periodontal disease. Furthermore, (I) is also useful for  
CC gut protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues, various immune deficiencies and  
CC disorders including severe combined immunodeficiency (SCID), bacterial or  
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,  
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
CC reactions and conditions, such as asthma or other respiratory problems.  
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,  
CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
CC analgesic effects or other pain reducing effects, immunoglobulin like  
CC activity and can act as an antigen in a vaccine composition to raise an  
CC immune response. AAU28020-AAU28395 represent novel human secreted protein  
CC amino acid sequences of the invention.  
XX  
SQ Sequence 166 AA;  
XX  
Query Match 36.3%; Score 97; DB 22; Length 166;  
Best Local Similarity 100.0%; Pred. No. 1e-82;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 75 DSVTSTFSKLRQLGPTVTOEFWDNLEKETEGLRQEMSKDLEEVKAKVOPYLDDFOKKQOE 134  
DB 61 dsvtstfsklrqlgptvtfwfnleketeglrqemskdleevkakvqpylddfqkkwqe 120  
OY 135 EMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEE 171  
DB 121 emelyrqkveplraelqegargarklhelqeklsplgee 157  
XX  
RESULT 16  
ID AAU30468 standard; Protein; 119 AA.  
XX  
AC AAU30468;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #959.  
XX

KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 297; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 119 AA;  
XX  
Query Match 34.5%; Score 92; DB 22; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3.7e-78;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 67 NLKLLDNMDSVTSTFSKLRQLGPTVTOEFWDNLEKETEGLRQEMSKDLEEVKAKVOPYLD 126  
DB 2 nlkllndwdsvtstfsklrqlgptvtfwfnleketeglrqemskdleevkakvpyld 61  
OY 127 DFQKKWQEMELYRQKVEPLRAELQEGAROKL 158  
DB 62 dfqkkwqeemelyrqkveplraelqegargarkl 93  
XX  
RESULT 17  
ID AAU30267 standard; Protein; 120 AA.  
XX  
AC AAU30267;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #758.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
KW



```
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US08656.
XX PR 18-APR-2000; 2000US-0552929.
XX PR 26-JAN-2001; 2001US-0770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy -
XX PS Claim 20; Page 270; 765pp; English.
XX CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX SQ Sequence 120 AA;

Query Match 25.5%; Score 68; DB 22; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 DSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLFEVKAKVQPYLDDFQKKWQE 134
Db 52 dsvtstfsklreglqpvtfqefwdnleketeglrqemskdleevkakvqpylddfqkkwqe 111

QY 135 EMELYRQK 142
Db 112 emelyrqk 119

RESULT 18
AAU30469
ID AAU30469 standard; Protein; 120 AA.
XX AC AAU30469;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #960.
XX KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PD
```

```
PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US08656.
XX PR 18-APR-2000; 2000US-0552929.
XX PR 26-JAN-2001; 2001US-0770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy -
XX PS Claim 20; Page 297; 765pp; English.
XX CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX SQ Sequence 120 AA;

Query Match 25.5%; Score 68; DB 22; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 DSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLFEVKAKVQPYLDDFQKKWQE 134
Db 52 dsvtstfsklreglqpvtfqefwdnleketeglrqemskdleevkakvqpylddfqkkwqe 111

QY 135 EMELYRQK 142
Db 112 emelyrqk 119

RESULT 19
AAU30268
ID AAU30268 standard; Protein; 318 AA.
XX AC AAU30268;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #759.
XX KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
```



```
XX 16-APR-2001; 2001WO-US08656.
PF
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 270; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 318 AA;

Query Match 18.0%; Score 48; DB 22; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 ATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALLEYTKKLTNQ 267
      |||||||
Db 246 atehlstlsekakpaledlrqglpvlsefkvsflsaleeytkkltntq 293

RESULT 20
AAU30470
ID AAU30470 standard; Protein; 359 AA.
XX
AC AAU30470;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #961.
XX
DE Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
```

```
XX PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 297; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 359 AA;

Query Match 18.0%; Score 48; DB 22; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.5e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 ATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALLEYTKKLTNQ 267
      |||||||
Db 312 atehlstlsekakpaledlrqglpvlsefkvsflsaleeytkkltntq 359

RESULT 21
AAR20164
ID AAR20164 standard; peptide; 42 AA.
XX
AC AAR20164;
XX
DT 01-APR-1992 (first entry)
XX
DE Apo AI polypeptide (A).
XX
DE Monoclonal antibody; MAB AI-11; epitope; diagnosis.
XX
OS Synthetic.
XX
FH Key
FH Peptide 1..28
FT Peptide 2..28
FT Peptide 7..28
FT Peptide 10..28
FT Peptide 11..28
FT Peptide 11..42
FT Peptide 13..28
XX
PN WO9118619-A.
XX
PD 12-DEC-1991.
XX
PF 07-JUN-1991; 91WO-US04038.
XX
PR 07-JUN-1990; 90US-0534761.
XX
PA (SCRI-) SCRIPPS CLINIC & RE.
XX
```

PI Curtiss LK, Banka CL, Bonnet DJ, Smith RS;  
XX  
DR WPI; 1992-007201/01.  
XX  
PT New Apo AI polypeptide(s) and anti-Apo antibodies - used to  
PT detect Apo AI in vascular fluid samples and increase  
PT LCAT-mediated cholesterol esterification in humans  
XX  
PS Claim 2; Page 68; 87pp; English.  
XX  
CC The peptides represented in AAR20164-65 are capable of immunologically  
CC mimicking an Apo AI epitope. They are useful in diagnosis and  
CC detection of Apo AI in fluids and for preparing anti-Apo AI  
CC antibodies. It is useful in therapeutic methods for increasing LCAT-  
CC mediated cholesterol esterification in humans.  
CC The Apo AI polypeptide is selected from the peptides indicated in  
CC the features and includes amino acids 13-28 defining a conserved  
CC native epitope on Apo AI capable of immunoreacting with monoclonal  
CC antibody MAB AI-11.  
XX  
SQ Sequence 42 AA;  
  
Query Match 15.7%; Score 42; DB 13; Length 42;  
Best Local Similarity 100.0%; Pred. No. 9e-32;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 108 QEMSKDLEEVKAKVQPYLDDEQKKWQEMELYRQKVEPLRAE 149  
Db 1 qemskdleevkavkqpylddfqkkwqemelyrqkveplrae 42  
|||||  
  
RESULT 22  
AAG62609  
ID AAG62609 standard; peptide; 26 AA.  
XX  
AC AAG62609;  
XX  
DT 06-SEP-2001 (first entry)  
XX  
DE Apolipoprotein fragment #3.  
XX  
KW Apolipoprotein; ApoA-1; atherosclerosis; coronary disease;  
KW cardiovascular disease; ischaemic heart disease; dyslipidaemia.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11 /label= OTHER  
FT /note= "optionally oxidised"  
XX  
PN WO200138395-A1.  
XX  
PD 31-MAY-2001.  
XX  
PF 27-NOV-2000; 2000WO-AU01463.  
XX  
PR 26-NOV-1999; 99AU-0004293.  
XX  
PA (HEAR-) HEART RES INST LTD.  
XX  
PI Stocker R, Wang XL, Wilcken D;  
XX  
DR WPI; 2001-355909/37.  
XX  
PT Novel oxidized form of apolipoprotein useful in diagnosis and treatment  
PT of diseases associated with oxidative stress such as cardiovascular  
PT diseases, in particular, atherosclerosis -  
XX  
PS Example 1; Page 28; 55pp; English.  
XX  
CC The present invention relates to oxidised apolipoprotein A-I (ApoA-I)

CC where at least Met residue 86 is oxidised to Met(O). This can be used in  
CC the prevention, diagnosis and treatment of lipid associated disorders,  
CC including coronary vascular disease, ischaemic heart disease,  
CC atherosclerosis and dyslipidaemias. The present sequence is a fragment of  
CC the ApoA-I protein isolated in the exemplification of the invention.  
XX  
SQ Sequence 26 AA;  
  
Query Match 9.7%; Score 26; DB 22; Length 26;  
Best Local Similarity 100.0%; Pred. No. 5.6e-17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 126 DDFQKKWQEMELYRQKVEPLRAEQ 151  
Db 1 ddfqkkwqemelyrqkveplraeq 26  
|||||  
  
RESULT 23  
AAU29834  
ID AAU29834 standard; Protein; 134 AA.  
XX  
AC AAU29834;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #325.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 199; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 134 AA;

Query Match 9.7%; Score 26; DB 22; Length 134;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 LFLTGSQARHFWQDEPPQSPWDRVK 36  
11 lfltsqarhfwqdeppqspwdrvk 36

RESULT 24  
AAO02278

ID AAO02278 standard; Protein; 151 AA.

AC AAO02278;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 16170.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR N-PSDB; AAI82209.

XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

PS Claim 20; SEQ ID NO 16170; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 151 AA;

Query Match 9.4%; Score 25; DB 22; Length 151;

Best Local Similarity 100.0%; Pred. No. 2.3e-15;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 118 KAKVQPYLDDFOKKWQEMELYRQK 142

Db 73 kakvqpylddfqkkwqemelyrpk 97

RESULT 25

AAAP92072

ID AAP92072 standard; peptide; 22 AA.

AC AAP92072;

DT 10-APR-1990 (first entry)

DE Apolipoprotein AI (Apo AI) polypeptide.  
KW Apolipoprotein AI; Apo AI; high-density lipoprotein; HDL;  
KW coronary artery disease; CAD; Apo AI epitope.

XX Key Location/Qualifiers

FT Peptide 12..22 /label= AI101-111

FT Peptide 4..12 /label= AI93-101

FT Peptide 1..22 /label= AI90-111

FT Peptide 11..16 /label= AI100-105

FT Peptide 7..12 /label= AI96-101

FT Peptide 1..22 /label= AI90-111

FT Peptide 1..16 /label= AI90-105

FT Peptide 6..16 /label= AI95-105

PN WO8904486-A.

PD 18-MAY-1989.

PF 02-NOV-1988; 88WO-US03903.

PR 02-NOV-1987; 87US-0116248.

XX (SCRI-) SCRIPPS CLINIC & RE.

PI Curtiss LK, Smith RR;

DR WPI; 1989-165740/22.

XX New monoclonal antibody and polypeptide antigens  
PT - directed against Apo AI-HDL epitope, useful in  
PT diagnostic assays

PS Claim 2; Table 1; Table 2; 62pp; English.

XX AI90-111 and smaller peptides contained within this sequence. AI95-105,  
CC AI190-105 and AI190-111 are specifically claimed. Such Apo AI peptides  
CC are capable of immunologically mimicking a native conserved Apo AI  
CC epitope. They may be helpful in the diagnosis of risk of coronary heart  
CC disease.

XX Sequence 22 AA;

Query Match 8.2%; Score 22; DB 10; Length 22;

Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 LEEVKAKVQPYLDDFOKKWQEE 135

Db 1 leevkakvqpylddfqkkwqee 22

RESULT 26  
AAG62608  
ID AAG62608 standard; peptide; 22 AA.  
XX  
AC AAG62608;  
XX  
DT 06-SEP-2001 (first entry)  
XX  
DE Apolipoprotein fragment #2.  
XX  
KW Apolipoprotein; ApoA-1; atherosclerosis; coronary disease;  
KM cardiovascular disease; ischaemic heart disease; dyslipidaemia.  
XX  
OS Unidentified.  
XX  
PN WO200138395-A1.  
XX  
PD 31-MAY-2001.  
XX  
PF 27-NOV-2000; 2000WO-AU01463.  
XX  
PR 26-NOV-1999; 99AU-0004293.  
XX  
PA (HEAR-) HEART RES INST LTD.  
XX  
PI Stocker R, Wang XL, Wilcken D;  
XX  
DR WPI; 2001-355909/37.  
XX  
PT Novel oxidized form of apolipoprotein useful in diagnosis and treatment  
PT of diseases associated with oxidative stress such as cardiovascular  
PT diseases, in particular, atherosclerosis -  
XX  
PS Example 1; Page 28; 55pp; English.  
XX  
CC The present invention relates to oxidised apolipoprotein A-I (ApoA-I)  
CC where at least Met residue 86 is oxidised to Met(O). This can be used in  
CC the prevention, diagnosis and treatment of lipid associated disorders,  
CC including coronary vascular disease, ischaemic heart disease,  
CC atherosclerosis and dyslipidaemias. The present sequence is a fragment of  
CC the ApoA-I protein isolated in the exemplification of the invention.  
XX  
SQ Sequence 22 AA;  
  
Query Match 8.2%; Score 22; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 152 EGAROKLHELQEKLSPLGEMR 173  
DB 1 egarqklhelqeklsplgeemr 22  
  
RESULT 27  
AAO12095  
ID AAO12095 standard; Protein; 154 AA.  
XX  
AC AAO12095;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 25987.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX

PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
DR N-PSDB; AAI92026.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 20; SEQ ID NO 25987; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 154 AA;  
  
Query Match 8.2%; Score 22; DB 22; Length 154;  
Best Local Similarity 100.0%; Pred. No. 1.5e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MKAAVLTAVLFLTGSQARHFW 22  
DB 9 mkaavltavflflltgsqarhfw 30  
  
RESULT 28  
AAP90956  
ID AAP90956 standard; peptide; 21 AA.  
XX  
AC AAP90956;  
XX  
DT 19-FEB-1990 (first entry)  
XX  
DE Apo AI epitope.  
XX  
KW High density lipoproteins.  
XX  
OS Synthetic.  
XX  
PN WO8909403-A.  
XX  
PD 05-OCT-1989.  
XX  
PF 27-MAR-1989; 89WO-US01262.  
XX  
PR 29-MAR-1988; 88US-0174698.  
XX  
PA (SCRI ) SCRIPPS CLINIC RES.  
XX  
PI Curtiss LK, Smith RS;  
XX





OY 111 SKDLEEVKAKVQPYLDDFQ 129  
DB 1 skdleevkakvqpylddfq 19

RESULT 31

AAV18676  
ID AAV18676 standard; Peptide; 18 AA.

AC AAV18676;

DT 09-JUL-1999 (first entry)

DE Peptide SEQ ID NO:270.

XX Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.  
XX Synthetic.

PN WO916409-A2.

PD 08-APR-1999.

PF 28-SEP-1998; 98WO-US20329.

PR 29-SEP-1997; 97US-0940136.

XX (BUTT/) BUTTNER K.  
PA (CORN/) CORNUT I.  
PA (DASS/) DASSEUX J.  
PA (DUFO/) DUFOURCQ J.  
PA (METZ/) METZ G.  
PA (SEKU/) SEKUL R.

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;  
PI Sekul R;

DR WPI; 1999-254921/21.

PT Nucleic acid encoding apolipoprotein A-I agonist peptides

PS Disclosure; Page 207; 232pp; English.

XX The present invention describes a nucleic acid (A) encoding an  
CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,  
CC which forms an amphipathic alpha-helix in presence of lipids. (A),  
CC optionally as a complex with lipids, and host cells that contain (A),  
CC are useful for gene therapy, or prevention, of diseases associated with  
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I  
CC deficiency, hypertriglyceridemia and metabolic syndrome, also to treat  
CC endotoxemia (septic shock). Host cells containing (A) can also be used  
CC to study the role of apoA-I in lipid metabolism. (B) can be used  
CC diagnostically, e.g. to measure serum HDL (particularly its  
CC subpopulation involved in retrograde cholesterol transport) and for  
CC imaging the circulatory system or HDL accumulations at fatty streaks.  
CC The present sequence represents a peptide from the present invention.

XX Sequence 18 AA;

Query Match 6.7%; Score 18; DB 20; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKA AVLFLAVLFLTGSQA 18  
ID 1 MKA AVLFLAVLFLTGSQA 18  
DB 1 mkaavltlavfltltsqa 18

RESULT 32

AAO08726  
ID AAO08726 standard; Protein; 38 AA.

AC AAO08726;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 22618.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.  
DR N-PSDB; AAI88657.

PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders

PS Claim 20; SEQ ID NO 22618; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 38 AA;

Query Match 6.7%; Score 18; DB 22; Length 38;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 250 KVSFLSALEYTKLNTQ 267  
ID 250 KVSFLSALEYTKLNTQ 267  
DB 21 kvsflsaleeytklntq 38

RESULT 33

AAAR98474  
ID AAR98474 standard; peptide; 17 AA.

AC AAR98474;

DT 25-FEB-1997 (first entry)

XX

DE Anti-arteriosclerotic peptide A.  
XX  
KW Anti-arteriosclerotic; human; apolipoprotein A-I; apoA-I;  
KW treatment; arteriosclerosis; prevention; restenosis;  
KW percutaneous transluminal coronary angioplasty; PTCA.  
XX  
OS Homo sapiens.  
XX  
PN JP08157492-A.  
XX  
PD 18-JUN-1996.  
XX  
PF 07-DEC-1994; 94JP-0331293.  
XX  
PR 07-DEC-1994; 94JP-0331293.  
XX  
PA (KAGA ) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.  
XX  
DR WPI; 1996-339187/34.  
XX  
PT Anti-arteriosclerotic peptide(s) derived from apo:lipoprotein A-I -  
PT used in the treatment of arteriosclerosis and re-stenosis  
XX  
PS Claim 3; Page 2; 7pp; Japanese.  
XX  
CC The present sequence is a specifically claimed  
CC anti-arteriosclerotic peptide derived from amino acid residues  
CC 227-243 of human apolipoprotein A-I, and its analogues. It may be  
CC used in the treatment of arteriosclerosis, and in the prevention  
CC of restenosis after percutaneous transluminal coronary  
CC angioplasty. An orally administered dose of 50 microg to 5 mg/kg  
CC of the peptide prevented arteriosclerosis in rabbits.  
XX  
SQ Sequence 17 AA;  
  
Query Match 6.4%; Score 17; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 251 VSFLSALEYTKKLNQ 267  
DB 1 vsflsaleeytkklnq 17  
  
RESULT 34  
AAR34033  
ID AAR34033 standard; Protein; 16 AA.  
XX  
AC AAR34033;  
XX  
DT 13-AUG-1993 (first entry)  
XX  
DE Sequence of apo AI from about residue 120 through about residue  
DE 135.  
XX  
KW Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.  
XX  
OS Synthetic.  
XX  
PN W09307165-A.  
XX  
PD 15-APR-1993.  
XX  
PF 09-OCT-1992; 92WO-US08634.  
XX  
PR 09-OCT-1991; 91US-0774633.  
PR 08-OCT-1992; 92US-0555555.  
PR 28-JUN-1992; 92US-0901706.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;

XX  
DR WPI; 1993-134378/16.  
XX  
PT Polypeptide mimic of native apo B-100 and native apo A-I - useful  
PT in assays for LDL and HDL in plasma samples  
XX  
PS Disclosure; Page 13 and page 35; 137pp; English.  
XX  
CC The inventors claim a portion of the polypeptide contg. apo B-100  
CC that immunoreacts with antibodies secreted by the hybridoma MB47  
CC having ATCC Accession No. 8746. Polypeptides specifically claimed  
CC include residues 217-297, 216-310, 216-331, 216-352, 216-377, 1-377,  
CC 205-297, 173-297, 140-297. DNA sequences encoding the polypeptides  
CC are also claimed. Also claimed are a fusion polypeptide that  
CC contains: (a) a first amino acid residue sequence up to 250 residues  
CC in length that includes residues 120-135 of apo A-I, (b) a second  
CC amino acid residue sequence up to 375 residues in length that  
CC includes residues 217-297 of apo B-100 and DNA encoding it.  
XX  
SQ Sequence 16 AA;

Query Match 6.0%; Score 16; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KVQPYLDPFQKKWQEE 135  
DB 1 kvqpyldpfqkkwqee 16

RESULT 35  
AAR72706  
ID AAR72706 standard; Peptide; 16 AA.  
XX  
AC AAR72706;  
XX  
DT 31-OCT-1995 (first entry)  
XX  
DE Human apo A-I fragment comprising residues 120-135.  
XX  
KW Apo A-I; LDL cholesterol; low density lipoprotein; lipid.  
XX  
OS Homo sapiens.  
XX  
PN US5408038-A.  
XX  
PD 18-APR-1995.  
XX  
PF 09-OCT-1991; 91US-0774633.  
XX  
PR 09-OCT-1991; 91US-0774633.  
PR 18-JUN-1992; 92US-0901706.  
PR 08-OCT-1992; 92US-0959946.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;  
XX  
DR WPI; 1995-161146/21.  
DR N-PSDB; AAQ89634.  
XX  
PT New apo:lipoprotein B-100 peptide(s) and fusion peptide(s) - used  
PT in assay systems for detecting LDL and HDL cholesterol levels in  
PT body fluids.  
XX  
PS Claim 10; Column 17; 41pp; English.  
XX  
CC A dispersible apo A-I/B-100 fusion polypeptide is claimed which  
CC contains a first AA sequence of apo A-I (see AAR72605) and that includes  
CC at least AA sequence positions 120-135 (see AAR72606). The two  
CC sequences are operatively linked. An exemplary linking sequence is  
CC AAR72707 whose encoding DNA can be ligated between an apo A-I and a

CC B-100 encoding DNA sequence.  
XX  
SQ Sequence 16 AA;

Query Match 6.0%; Score 16; DB 16; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KVGPLYLDDFQKKWQOE 135  
| | | | | | | | | | | | | | | | | |  
Db 1 kvqpylddfqkkwqee 16

## RESULT 36

AAV27065  
ID AAV27065 standard; peptide; 16 AA.

XX  
AC AAV27065;

XX  
DT 18-OCT-1999 (first entry)

XX  
DE Peptide Seq ID No: 1 of JP11209399.

XX  
KW Medicinal composition; lipoprotein; megakaryocytic; antithrombin III;  
KW alveolar forming activity; C1 inhibitor; human; thrombocytopenia;  
KW platelet.

XX  
OS Synthetic.

XX  
PN JP11209399-A.

XX  
PD 03-AUG-1999.

XX  
PF 28-MAY-1998; 98JP-0146860.

XX  
PR 20-NOV-1997; 97JP-0319587.

XX  
PA (ASAH ) ASAH KASEI KOGYO KK.

XX  
DR WPI; 1999-48816/41.

XX  
PT New medicinal composition composed of an effective ingredient of  
PT heavy specific gravity lipoprotein - useful for effective production  
PT of platelets and treatment of thrombocytopenia

XX  
PS Disclosure; Page 12; 14pp; Japanese.

XX  
CC The invention provides a new medicinal composition composed of an  
CC effective ingredient of a heavy specific gravity lipoprotein containing  
CC megakaryocytic projected alveolar forming activity, particularly non-  
CC absorptive in zinc chelating gel, especially containing antithrombin III,  
CC having C1 inhibitor activity. The lipoprotein is prepared by: (a) heating  
CC human blood at 60 deg. C for 10 minutes; (b) removal of protein to give a  
CC supernatant; (c) dialysis of the resultant supernatant against 10 mM  
CC NaOAc buffer at pH 4.5, to give a supernatant; (d) dialysis of the  
CC resultant supernatant against 10 mM HEPES buffer containing 150 mM NaCl  
CC at pH 7.4, to give a supernatant; (e) gel filtration with 10 mM HEPES  
CC buffer containing 150 mM NaCl at pH 7.4 to give a fraction having the  
CC activity; (f) contacting the active fraction with a zinc chelating gel  
CC under 10 mM HEPES buffer containing 150 mM NaCl at pH 7.4; (g) collection  
CC of non-absorptive fraction; (h) ultracentrifugation of the fraction  
CC adjusted to S.G. 1.21; and (i) collection of the floating yellow fraction  
CC to give the aimed heavy specific gravity lipoprotein containing  
CC megakaryocytic projected alveolar forming activity; used for treatment of  
CC thrombocytopenia. The new composition is useful for effective production  
CC of platelets and treatment of thrombocytopenia.

XX  
SQ Sequence 16 AA;

Query Match 6.0%; Score 16; DB 20; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.2e-08;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DEPPQSPWDRVKDLAT 40  
| | | | | | | | | | | | | | | | | |  
Db 1 deppqspwdrvkdlat 16

## RESULT 37

AAG62607  
ID AAG62607 standard; peptide; 16 AA.

XX  
AC AAG62607;

XX  
DT 06-SEP-2001 (first entry)

XX  
DE Apolipoprotein fragment #1.

XX  
KW Apolipoprotein; ApoA-1; atherosclerosis; coronary disease;  
KW cardiovascular disease; ischaemic heart disease; dyslipidaemia.

XX  
OS Unidentified.

XX  
FH Key Location/Qualifiers

FT Modified-site 14

FT FT /label= OTHER  
FT /note= "optionally oxidised"

XX  
PN WO200138395-A1.

XX  
PD 31-MAY-2001.

XX  
PF 27-NOV-2000; 2000WO-AU01463.

XX  
PR 26-NOV-1999; 99AU-0004293.

XX  
PA (HEAR-) HEART RES INST LTD.

XX  
PI Stocker R, Wang XL, Wilcken D;

XX  
DR WPI; 2001-355909/37.

XX  
PT Novel oxidized form of apolipoprotein useful in diagnosis and treatment  
PT of diseases associated with oxidative stress such as cardiovascular  
PT diseases, in particular, atherosclerosis

XX  
PS Example 1; Page 28; 55pp; English.

XX  
CC The present invention relates to oxidised apolipoprotein A-I (ApoA-I)  
CC where at least Met residue 86 is oxidised to Met(O). This can be used in  
CC the prevention, diagnosis and treatment of lipid associated disorders,  
CC including coronary vascular disease, ischaemic heart disease,  
CC atherosclerosis and dyslipidaemias. The present sequence is a fragment of  
CC the ApoA-I protein isolated in the exemplification of the invention.

XX  
SQ Sequence 16 AA;

Query Match 6.0%; Score 16; DB 22; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 DNEKETEGLRQEMSK 112  
| | | | | | | | | | | | | | | | | |  
Db 1 dneketeglrgemsk 16

## RESULT 38

AAP60986  
ID AAP60986 standard; Peptide; 16 AA.

XX  
AC AAP60986;

XX  
DT 08-AUG-1991 (first entry)

```
XX DE Sequence of immunogenic fragment of apolipoprotein (APL) ApoA1.
XX KW Apolipoprotein; immunoassay; antibody.
XX OS Homo sapiens.
XX PN W08604144-A.
XX PD 17-JUL-1986.
XX PF 26-DEC-1985; 85WO-US02569.
XX PR 31-DEC-1984; 84US-0688040.
XX PR 26-DEC-1985; 85US-0905584.
XX PA (ITGE-) INT GENETIC ENG INC.
XX PI Fareed G, Sen A;
XX DR WPI; 1986-196930/30.
XX PT Peptide fragments of human apolipoprotein - used for producing
XX PT type-specific antibodies for immunoassay
XX PS Claim 19; page 38; 53pp; English.
XX CC The peptides of the invention are conjugated with carrier proteins
XX CC and used to produce type-specific, non-cross-reactive antibodies by
XX CC immunisation. The antibodies may then be used in immunoassays to
XX CC identify and quantitate specific APLs. Peptide fragments without
XX CC the C-terminal Cys residue are also claimed.
SQ Sequence 16 AA;

OY 184 RTHLAPYSDELRLQRL 198
   |||||
Db 1 rthlapysdelrgrl 15

Query Match          5.6%; Score 15; DB 7; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 39
AAP60985
ID AAP60985 standard; Protein; 14 AA.
XX AC AAP60985;
XX DT 08-AUG-1991 (first entry)
XX DE Sequence of immunogenic fragment of apolipoprotein (APL) ApoA1.
XX KW Apolipoprotein; immunoassay; antibody.
XX OS Homo sapiens.
XX PN W08604144-A.
XX PD 17-JUL-1986.
XX PF 26-DEC-1985; 85WO-US02569.
XX PR 31-DEC-1984; 84US-0688040.
XX PR 26-DEC-1985; 85US-0905584.
XX PA (ITGE-) INT GENETIC ENG INC.
XX PI Fareed G, Sen A;
XX DR WPI; 1986-196930/30.
```

```
XX PT Peptide fragments of human apolipoprotein - used for producing
XX PT type-specific antibodies for immunoassay
XX PS Claim 18; page 37; 53pp; English.
XX CC The peptides of the invention are conjugated with carrier proteins
XX CC and used to produce type-specific, non-cross-reactive antibodies by
XX CC immunisation. The antibodies may then be used in immunoassays to
XX CC identify and quantitate specific APLs. Peptide fragments without
XX CC the C-terminal Cys residue are also claimed.
SQ Sequence 14 AA;

OY 25 DEPPQSPWDRVKD 37
   |||||
Db 1 deppqspwdrvk 13

Query Match          4.9%; Score 13; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 40
AAP92076
ID AAP92076 standard; peptide; 12 AA.
XX AC AAP92076;
XX DT 10-APR-1990 (first entry)
XX DE Apolipoprotein AI (Apo AI) polypeptide.
XX KW Apolipoprotein AI; Apo AI; high-density lipoprotein; HDL;
XX KW coronary artery disease; CAD; Apo AI epitope; AI105-116.
XX PN W08904486-A.
XX PD 18-MAY-1989.
XX PF 02-NOV-1988; 88WO-US03903.
XX PR 02-NOV-1987; 87US-0116248.
XX PA (SCRI-) SCRIPPS CLINIC & RE.
XX PI Curtiss LK, Smith RR;
XX DR WPI; 1989-165740/22.
XX DT New monoclonal antibody and polypeptide antigens
XX DT - directed against Apo AI-HDL epitope, useful in
XX DT diagnostic assays
XX PS Page 31; ; 62pp; English.
XX CC An Apo AI peptide capable of immunologically mimicking a native
XX CC conserved Apo AI epitope. It may be helpful in the diagnosis of risk of
XX CC coronary heart disease.
SQ Sequence 12 AA;

OY 129 QKKWQEMELLYR 140
   |||||
Db 1 qkkwqeemelyr 12

Query Match          4.5%; Score 12; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 41
AAV41961
ID AAY41961 standard; Peptide; 12 AA.
XX
AC AAY41961;
XX
DT 09-DEC-1999 (first entry)
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #112.
XX
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
OS Homo sapiens.
XX
PN WO9947925-A2.
XX
PD 23-SEP-1999.
XX
PF 15-MAR-1999; 99WO-GB00763.
XX
PR 13-MAR-1998; 98GB-0005477.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Parekh RB, Patel TP, Townsend RR;
XX
DR WPI; 1999-571871/48.
XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis -
XX
PS Disclosure; Page 20; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a
CC two-dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC AAY42103 represent expression reference protein isoform peptides and
CC AA225066 to AA225068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 4.5%; Score 12; DB 20; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 VSFLSALIEYTK 262
| | | | | | | | | |
Db 1 vsflsalseeytk 12
```

RESULT 42

```
AAV42027
ID AAY42027 standard; Peptide; 12 AA.
XX
AC AAY42027;
XX
DT 09-DEC-1999 (first entry)
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #178.
XX
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
OS Homo sapiens.
XX
PN WO9947925-A2.
XX
PD 23-SEP-1999.
XX
PF 15-MAR-1999; 99WO-GB00763.
XX
PR 13-MAR-1998; 98GB-0005477.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Parekh RB, Patel TP, Townsend RR;
XX
DR WPI; 1999-571871/48.
XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis -
XX
PS Disclosure; Page 21; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a
CC two-dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AAY41844 to AAY42100 represent RPI peptides, AAY42101 to
CC AAY42103 represent expression reference protein isoform peptides and
CC AA225066 to AA225068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 4.5%; Score 12; DB 20; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 VSFLSALIEYTK 262
| | | | | | | | | |
Db 1 vsflsalseeytk 12
```

RESULT 43  
AAB87196



ID AAB87196 standard; Peptide; 12 AA.  
XX  
AC AAB87196;  
XX  
DT 10-MAY-2001 (first entry)  
XX  
DE Breast-cancer associated protein isoform BPI-9 peptide #4.  
XX  
KW Human; breast cancer; breast cancer associated protein isoform; BPI;  
KW breast cancer associated feature; BF; diagnosis; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200113117-A2.  
XX  
PD 22-FEB-2001.  
XX  
PF 14-AUG-2000; 2000WO-GB03143.  
XX  
PR 13-AUG-1999; 99GB-0019258.  
PR 30-MAR-2000; 2000GB-0007754.  
XX  
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX  
PI Herath HMAc;  
XX  
DR WPI; 2001-211252/21.  
XX  
PT Screening, diagnosis or prognosis of breast cancer, by analyzing a  
PT sample of serum or plasma by two dimensional electrophoresis to detect  
PT the presence or level of a breast cancer-associated feature -  
XX  
PS Disclosure; Page 23; 146pp; English.  
XX  
CC The present invention describes a method for the screening, diagnosis or  
CC prognosis of breast cancer (BC), determining the stage or severity of BC,  
CC and monitoring the effect of therapy administered to a subject having BC,  
CC comprising analysing a sample of body fluid by two dimensional  
CC electrophoresis to generate a two-dimensional array of features,  
CC comprising a chosen feature whose abundance correlates with BC or  
CC predicts the onset or course of BC. The method (I) involves:  
CC (a) analysing a sample of body fluid from the subject by two-dimensional  
CC electrophoresis to generate a two-dimensional array of features,  
CC comprising a chosen feature whose relative abundance correlates with BC  
CC or predicts the onset of BC; and (b) comparing the abundance of each  
CC chosen feature in the sample with the abundance of that chosen feature  
CC in the body fluid from one or more persons free from BC, or with a  
CC previously determined reference range for that feature in subjects free  
CC from BC, or with the abundance of an expression reference feature (ERF)  
CC in the test sample. The method is useful for screening, diagnosis or  
CC prognosis of breast cancer, determining the stage or severity of BC,  
CC monitoring the effect of therapy administered to a subject having BC,  
CC and for identifying a subject at risk of developing BC. AAB87186 to  
CC AAB87340 represents breast cancer associated protein isoform (BPI)  
CC peptide sequences, and AAF91643 to AAF91848 represent BPI probes used in  
CC the exemplification of the present invention.  
XX  
SQ Sequence 12 AA;  
XX  
Query Match 4.5%; Score 12; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00035;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 VSFLSALEEYTK 262  
Db 1 vsflsaleeytk 12

RESULT 44  
AAP60987  
ID AAP80987 standard; Peptide; 13 AA.  
XX

AC AAP60987;  
XX  
DT 08-AUG-1991 (first entry)  
XX  
DE Sequence of immunogenic fragment of apolipoprotein (APL) ApoE1.  
XX  
KW Apolipoprotein; immunoassay; antibody.  
XX  
OS Homo sapiens.  
XX  
PN WO8604144-A.  
XX  
PD 17-JUL-1986.  
XX  
PF 26-DEC-1985; 85WO-US02569.  
XX  
PR 31-DEC-1984; 84US-0688040.  
PR 26-DEC-1985; 85US-0905584.  
XX  
PA (ITGE-) INT GENETIC ENG INC.  
XX  
PI Fareed G, Sen A;  
XX  
DR WPI; 1986-196930/30.  
XX  
PT Peptide fragments of human apo:lipoprotein - used for producing  
PT type-specific antibodies for immunoassay  
XX  
PS Claim 20; page 38; 53pp; English.  
XX  
CC The peptides of the invention are conjugated with carrier proteins  
CC and used to produce type-specific, non-cross-reactive antibodies by  
CC immunisation. The antibodies may then be used in immunoassays to  
CC identify and quantitate specific APLs. Peptide fragments without  
CC the C-terminal Cys residue are also claimed.  
XX  
SQ Sequence 13 AA;  
XX  
Query Match 4.5%; Score 12; DB 7; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00038;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 VKAKVQPYLDDF 128  
Db 1 vkakvpylddf 12

RESULT 45  
AAB87236  
ID AAB87236 standard; Peptide; 22 AA.  
XX  
AC AAB87236;  
XX  
DT 10-MAY-2001 (first entry)  
XX  
DE Breast-cancer associated protein isoform BPI-28 peptide #4.  
XX  
KW Human; breast cancer; breast cancer associated protein isoform; BPI;  
KW breast cancer associated feature; BF; diagnosis; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200113117-A2.  
XX  
PD 22-FEB-2001.  
XX  
PF 14-AUG-2000; 2000WO-GB03143.  
XX  
PR 13-AUG-1999; 99GB-0019258.  
PR 30-MAR-2000; 2000GB-0007754.  
XX  
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX



XX WPI; 1999-571871/48.  
XX  
XX  
PT diagnosis of human rheumatoid arthritis by two-dimensional  
PT electrophoresis -  
XX  
XX  
PS Disclosure; Page 21; 157pp; English.  
XX  
XX A method has been developed for the diagnosis of human rheumatoid  
CC arthritis (RA) using two-dimensional electrophoresis to generate a  
CC two-dimensional array of features. The method can be used for screening,  
CC diagnosis and prognosis of RA in a subject or for monitoring the effect  
CC of an anti-RA drug or therapy administered to a subject. The method  
CC comprises: (a) analysing a sample of serum or plasma and optionally  
CC synovial fluid by two-dimensional electrophoresis, to generate a two-  
CC dimensional array of features; (b) identifying at least one chosen  
CC feature whose relative abundance correlates with the presence or absence  
CC of RA; and (c) comparing the abundance of each chosen feature in the  
CC sample with the abundance of that chosen feature in serum or plasma from  
CC one or more persons without RA, where the relative abundance of the  
CC chosen feature or features in the sample indicates the presence or  
CC absence of RA in the subject. The method can also be used in clinical  
CC studies for testing drugs for therapy of RA, for purification of RA-  
CC diagnostic protein isoforms (RPIs), and for production of antibodies to  
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify  
CC compounds that promote or inhibit their activity, which are then used as  
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy  
CC protocols. AAy41844 to AAy42100 represent RPI peptides, AAy42101 to  
CC AAy42103 represent expression reference protein isoform peptides and  
CC AAz25066 to AAz25068 represent degenerate probes for RPIs, which are all  
CC used in the exemplification of the present invention.  
XX  
SO Sequence 11 AA;

Query Match	4.1%;	Score 11;	DB 20;	length 11;
Best Local Similarity	100.0%;	Pred. No. 0.0028;		
Matches	11;	Conservative	0;	Mismatches 0;
Indels				0;
Gaps				0;
QY	37	DLATVYVDVLK	47	
Db	1	dlatvyvdvdk	11	

RESULT	48
ABB56060	
ID	ABB56060 standard; peptide; 11 AA.
XX	
AC	ABB56060;
XX	
DT	15-FEB-2002 (first entry)
XX	
DE	Vascular dementia-associated protein isoform (VPI) 260.
XX	
KW	Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW	diagnosis; prognosis; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	WO200169261-A2.
XX	
PD	20-SEP-2001.
XX	
PF	14-MAR-2001; 2001WO-GB01106.
XX	
PR	15-MAR-2000; 2000GB-0006285.
PR	24-NOV-2000; 2000GB-0028734.
PR	28-NOV-2000; 2000US-0724391.
XX	
PA	(OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX	
PI	Heíath HMAC, Parekh RB, Rohlf C;
XX	

DR WPI, 2001-557937/62.

XX Screening, diagnosis or prognosis of vascular dementia (VD), useful for  
PT determining stage of VD and monitoring the effect of VD therapy,  
PT comprises analysing body fluid by 2-dimensional electrophoresis for  
PT features correlated with VD -

XX  
PS  
PS Claim 6; Page 35; 151pp; English.

XX  
XX  
CC The invention relates to screening, diagnosis or prognosis of Vascular  
CC Dementia (VD) in a subject comprising analysing body fluid from the  
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of  
CC features containing at least one chosen feature whose relative abundance  
CC correlates with the presence, absence, stage or severity of VD or  
CC predicts the onset or course of VD, especially detecting in a sample of  
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated  
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the  
CC specification. Detecting VD-associated features and VPI is useful for the  
CC screening, diagnosis or prognosis of VD, for determining the stage or  
CC severity of VD, for identifying a subject at risk of VD or for  
CC monitoring the effect of therapy administered to a subject having VD.  
CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are  
CC useful for the treatment of VD and for gene therapy.

Query Match	4.1%;	Score 11;	DB 22;	length 11;
Best Local Similarity	100.0%;	Pred. No. 0.0028;		
Matches 11; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	185	THLAPYSDELR	195
Db	1	thlapydselr	11

RESULT	49
AAU24926	
ID	AAU24926 standard; Peptide; 11 AA.
XX	
AC	AAU24926;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Schizophrenia-Associated Protein Isoform (SPI) peptide #155.
XX	
KW	Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
KW	neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX	
OS	Homo sapiens.
XX	
PN	WO200162785-A2.
XX	
PD	30-AUG-2001.
XX	
PF	23-FEB-2001; 2001WO-GB00792.
XX	
PR	24-FEB-2000; 2000GB-0004A15.
PR	28-NOV-2000; 2000US-0750395.
XX	
PA	(OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX	
PI	Herath HMAc, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
XX	
DR	WPI; 2001-570624/64.
XX	
PT	New schizophrenia associated protein isoforms and encoding nucleic acid
PT	molecules, useful for treatment, diagnosis and prognosis of
PT	schizophrenia and screening for potential drugs for treatment and new
PT	drug targets -
XX	
PS	Disclosure; Page 32; 148pp; English.
XX	

CC The sequence represents a schizophrenia-associated protein isoform (SPI).  
 CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable  
 CC in cerebrospinal fluid, serum or plasma and are useful markers of  
 CC schizophrenia. The sequences can be used for treatment and diagnosis of  
 CC schizophrenia, screening, prognosis, monitoring the results of therapy,  
 CC identifying patients most likely to respond to a particular therapy and  
 CC identification of new targets for drug treatment. SPI DNA is useful as a  
 CC nucleic acid probe to detect the presence of nucleic acids or SPIS.  
 XX  
 SQ Sequence 11 AA;

Query Match 4.1%; Score 11; DB 22; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0028;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 185 THLAPYSDELR 195  
 Db 1 thlapysdelr 11

RESULT 50  
 AAU15270  
 ID AAU15270 standard; Peptide: 11 AA.

XX AAU15270;  
 DT 24-OCT-2001 (first entry)

XX Schizophrenia-associated isoform peptide #155.

XX Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;  
 KW neurological disorder; neuropathy.

XX Homo sapiens.

XX WO200163293-A2.

XX 30-AUG-2001.

XX 23-FEB-2001; 2001WO-GB00783.

XX 24-FEB-2000; 2000GB-0004415.

XX 28-NOV-2000; 2000US-0750395.

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX Herath HMC, Parekh RB, Rohlf C;

XX WPI; 2001-502868/55.

PT Diagnosing and monitoring Schizophrenia by detecting the presence of  
 PT Schizophrenia Associated Features and Schizophrenia Associated Protein  
 PT Isoforms in samples of cerebrospinal fluid -

PS Claim 6; Page 32; 160pp; English.

XX The invention relates to methods and compositions for screening,  
 CC diagnosis and prognosis of Schizophrenia. The method involves detecting  
 CC the presence of Schizophrenia (SCH) Associated Features (SFs) and SCH  
 CC Associated Protein Isoforms (SPIS) in samples, e.g. by electrophoresis,  
 CC immunoassay or hybridisation assay, for diagnosing and monitoring SCH,  
 CC studying the effectiveness of treatments and for identifying potential  
 CC therapeutic agents. The method is used for (1) screening or diagnosis of  
 CC SCH and the relative abundance of at least 1 chosen feature or diagnosis of  
 CC with the presence or absence of SCH; and (2) monitoring the effect of  
 CC therapy administered to a subject with SCH and the relative abundance of  
 CC at least 1 chosen feature which correlates with the severity of SCH.  
 CC The expression and activity of the SFs, SPIS and related molecules  
 CC (e.g. secondary messengers) are studied to diagnose SCH, monitor the  
 CC progress of the disorder and the effectiveness of treatment and as  
 CC targets to identify and produce potential therapeutic agents for the  
 CC treatment of SCH. The paucity of detectable neuronal defects

CC distinguishes neuropsychiatric disorders such as SCH from neurological  
 CC disorders, where manifestations of anatomical and biochemical changes  
 CC have been identified in many cases. Consequently the identification and  
 CC characterisation of cellular and/or molecular causative defects and  
 CC neuropathies are necessary for improved treatment of neuropsychiatric  
 CC disorders. AAU15114-AAU15762 represent the amino acid sequences of  
 CC schizophrenia-associated isoforms used in the method of the invention.  
 XX  
 SQ Sequence 11 AA;

Query Match 4.1%; Score 11; DB 22; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0028;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 185 THLAPYSDELR 195  
 Db 1 thlapysdelr 11

Search completed: September 22, 2002, 12:22:53  
 Job time: 978 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 22, 2002, 12:01:05 ; Search time 66.91 Seconds  
(without alignments)  
383.438 Million cell updates/sec

Title: US-09-803-918A-2

Perfect score: 1362

Sequence: 1 MKAAVLTLAVLFLTGSQARH.....SFKVSEFLSALLEYTKKLNTQ 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summariesDatabase : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	267	1 LPHUA1	apolipoprotein A-I
2	1306	95.9	267	2 JS0079	apolipoprotein A-I
3	1299	95.4	267	1 A26529	apolipoprotein A-I
4	1161.5	85.3	266	1 LPDGA1	apolipoprotein A-I
5	1100.5	80.8	266	1 LPRB12	apolipoprotein A-I
6	1100	80.8	265	2 A46018	apolipoprotein A-I
7	1093	80.2	265	2 JT0672	apolipoprotein A-I
8	1087.5	79.8	264	2 S31394	apolipoprotein A-I
9	1082	79.4	265	1 LPRB1B	apolipoprotein A-I
10	1082	79.4	265	2 A56858	apolipoprotein A-I
11	974.5	71.5	241	2 A24998	apolipoprotein A-I
12	924	67.8	231	2 JQ0704	apolipoprotein A-I
13	904.5	66.4	264	2 S22420	apolipoprotein A-I
14	898.5	66.0	262	2 JC1237	apolipoprotein A-I
15	849.5	62.4	259	2 A24700	apolipoprotein A-I
16	688.5	50.6	264	1 LPCHA1	apolipoprotein A-I
17	663.5	48.7	264	2 JC5456	apolipoprotein A-I
18	644.5	47.3	246	2 A61448	apolipoprotein A-I
19	617.5	45.3	164	2 S21830	apolipoprotein A-I
20	291	21.4	429	2 S29565	apolipoprotein A-I
21	278.5	20.4	396	1 LPHUA4	apolipoprotein A-I
22	267	19.6	391	1 LPRTA4	apolipoprotein A-I
23	253	18.6	399	2 C40892	apolipoprotein A-I
24	251	18.4	391	2 B40892	apolipoprotein A-I
25	251	18.4	395	2 A40892	apolipoprotein A-I
26	249.5	18.3	401	2 A47141	apolipoprotein A-I
27	238.5	17.5	394	2 A25281	apolipoprotein A-I
28	225.5	16.6	258	2 JH0472	apolipoprotein A-I
29	209	15.3	311	2 JU0036	apolipoprotein E p

30	198	14.5	312	1 LPRTF	apolipoprotein E p
31	194.5	14.3	1547	2 T28657	blackjack protein,
32	190.5	14.0	317	2 A28792	apolipoprotein E p
33	189.5	13.9	317	2 S03185	apolipoprotein E p
34	188	13.8	317	2 S33450	apolipoprotein E -
35	184	13.5	298	2 S12635	apolipoprotein E p
36	182	13.4	329	2 JC5566	apolipoprotein E p
37	178.5	13.1	291	2 C60940	apolipoprotein E -
38	174.5	12.8	317	1 LPHUE	apolipoprotein E p
39	166.5	12.2	316	2 JC6549	apolipoprotein E p
40	160	11.7	316	2 S26478	apolipoprotein E -
41	160	11.7	316	2 I45996	apolipoprotein E -
42	147	10.8	311	2 A45951	apolipoprotein E p
43	146	10.7	513	2 S08381	keratin, 58K type
44	143.5	10.5	771	1 A33430	h-caldesmon - chic
45	137	10.1	470	2 T23512	hypothetical prote

## ALIGNMENTS

RESULT 1  
LPHUA1  
apolipoprotein A-I precursor [validated] - human  
N/Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C/Species: Homo sapiens (man)  
C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000  
C/Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010;  
6197  
R;Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A/Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A/Reference number: A90947; MUID:85026665  
A/Accession: A90947  
A/Molecule type: DNA  
A/Residues: 1-267 <SEI>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
A/Accession: B90947  
A/Molecule type: mRNA  
A/Residues: 1-267 <SE2>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R;Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988  
A/Title: Sequence and expression of Tangier apoA-I gene.  
A/Reference number: S02373; MUID:88196137  
A/Accession: S02373  
A/Molecule type: DNA  
A/Residues: 1-267 <MAK>  
A/Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
R;Shoulders, C.C.; Kornblith, A.R.; Munro, B.S.; Baralle, F.E.  
Nucleic Acids Res. 11, 2827-2837, 1983  
A/Title: Gene structure of human apolipoprotein A1.  
A/Reference number: A93465; MUID:83220822  
A/Accession: A93465  
A/Molecule type: DNA  
A/Residues: 1-267 <SHO>  
A/Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518;  
R;Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
A/Title: Isolation and characterization of the human apolipoprotein A-I gene.  
A/Reference number: A21147; MUID:84016011  
A/Accession: A21147  
A/Molecule type: DNA  
A/Residues: 1-267 <KAR>  
A/Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.  
Nucleic Acids Res. 12, 3917-3932, 1984  
A/Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundan  
A/Reference number: A93519; MUID:84221405  
A/Accession: A93519  
A/Molecule type: mRNA  
A/Residues: 1-267 <SHA>  
A/Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753

A:Accession: B93519  
 A:Molecule type: DNA  
 A:Residues: 1-24 <SH2>  
 R:Cheung, P.; Chan, L.  
 Nucleic Acids Res. 11, 3703-3715, 1983  
 A:Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
 A:Reference number: A93472; MUID:83220772  
 A:Accession: A93472  
 A:Molecule type: mRNA  
 A:Residues: 1-267 <CHE>  
 A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R:Law, S.W.; Brewer Jr., H.B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
 A:Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
 A:Reference number: A94010; MUID:84119464  
 A:Accession: A94010  
 A:Molecule type: mRNA  
 A:Residues: 1-267 <LAW>  
 A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R:Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
 A:Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
 A:Reference number: A21118; MUID:83195100  
 A:Accession: A21118  
 A:Molecule type: protein  
 A:Residues: 1-24 <ZAN>  
 R:Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
 Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
 A:Title: Human plasma proapoA-I: isolation and amino-terminal sequence.  
 A:Reference number: A90112; MUID:83256553  
 A:Accession: A90112  
 A:Molecule type: protein  
 A:Residues: 19-27 <BRE>  
 R:Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
 Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
 A:Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high density lipoproteins  
 A:Reference number: A90209; MUID:78123731  
 A:Accession: A90209  
 A:Molecule type: protein  
 A:Residues: 25-57, 'Q', 59-169, 'QQ', 172-267 <BR2>  
 R:Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
 J. Clin. Invest. 82, 803-807, 1988  
 A:Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A-I)  
 A:Reference number: A30516; MUID:88331387  
 A:Accession: A30516  
 A:Molecule type: protein  
 A:Residues: 25-56 <YUI>  
 R:Nichols, W.C.; Dwulet, F.E.; Liepnieks, J.; Benson, M.D.  
 Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
 A:Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid.  
 A:Reference number: A31582; MUID:89050104  
 A:Accession: A31582  
 A:Molecule type: protein  
 A:Residues: 25-49, 'R', 51-85, 'D', 87-107 <NIC>  
 A:Note: variant sequence from patient with familial amyloidotic polyneuropathy type III  
 R:Manjunath, P.; Marcel, Y.L.; Uma, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A.  
 J. Biol. Chem. 264, 16853-16857, 1989  
 A:Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
 A:Reference number: A34409; MUID:89380318  
 A:Accession: A34409  
 A:Molecule type: protein  
 A:Residues: 25-48 <MAN>  
 R:Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
 A:Title: Structural requirements of human preproapolipoprotein AI for translocation and secretion  
 A:Reference number: S02737; MUID:89149957  
 A:Accession: S02737  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-6, 'AV', 9, 'LV', 12-29 <STO>  
 A:Note: Part of this sequence, including the amino end of the mature protein, was confirmed by Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A:Title: Transient expression of wild type and mutant human apolipoprotein AI in COS cells  
 A:Reference number: S16197; MUID:92029676  
 A:Contents: annotation; extension of studies in reference S02737  
 R:Stoffel, W.; Kruger, E.; Deutzmann, R.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
 A:Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processed in vitro  
 A:Reference number: A19913; MUID:83236195  
 A:Accession: B19913  
 A:Molecule type: protein  
 A:Residues: 1-6, 'X', 8-13, 'XXX', 17-18, 'XX', 21, 'X', 23-25, 'X', 27-29 <ST2>  
 R:Elmholm, C.; Boxas, S.E.; Tenkanen, H.; Kirsbaum, L.; Metso, J.; Murphy, B.; Walke, B.  
 Biochim. Biophys. Acta 1086, 255-260, 1991  
 A:Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
 A:Reference number: A56815; MUID:92075698  
 A:Accession: A56815  
 A:Molecule type: protein  
 A:Residues: 25-31, 'P', 33 <EHN>  
 A:Experimental source: serum  
 A:Note: sequence extracted from NCBI backbone (NCBI:69759)  
 A:Note: 32-Trp was also found  
 R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
 Biochemistry 33, 1988-1993, 1994  
 A:Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins  
 A:Reference number: A54223; MUID:94162201  
 A:Accession: A54223  
 A:Molecule type: protein  
 A:Residues: 25-39 <KUN>  
 R:Mogulievsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.; DNA 8, 429-436, 1989  
 A:Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: purification and characterization  
 A:Reference number: I39476; MUID:89377481  
 A:Accession: I39476  
 A:Molecule type: mRNA  
 A:Residues: 19-267 <RES>  
 A:Cross-references: GB:M29068; NID:g178774; PIDN:AAA51747.1; PID:g178775  
 R:Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B.  
 J. Biol. Chem. 263, 18530-18536, 1988  
 A:Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by thyroid hormone  
 A:Reference number: I39475; MUID:89054040  
 A:Accession: I39475  
 A:Molecule type: DNA  
 A:Residues: 1-14 <RE2>  
 A:Cross-references: GB:J04066; NID:g178763; PIDN:AAA51746.1; PID:g553183  
 R:Breslow, J.L.  
 Annu. Rev. Biochem. 54, 699-727, 1985  
 A:Title: Human apolipoprotein molecular biology and genetic variation.  
 A:Reference number: A90042; MUID:85278004  
 A:Contents: annotation; review of sequences, variants and gene location  
 R:Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B.  
 J. Biol. Chem. 261, 3911-3914, 1986  
 A:Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation  
 A:Reference number: A92577; MUID:86140194  
 A:Contents: annotation; acylation with palmitate  
 A:Note: an undetermined serine or threonine is acylated by fatty acid; the acylating agent is not determined  
 R:Law, S.W.; Brewer, H.B.  
 J. Biol. Chem. 260, 12810-12814, 1985  
 A:Title: Tangier disease: The complete mRNA sequence encoding for preproapo-A-I.  
 A:Reference number: I55236; MUID:86008382  
 A:Accession: I55236  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-143, 'D', 145-267 <RE3>  
 A:Cross-references: GB:M11791; NID:g178776; PIDN:AAA35545.1; PID:g178777  
 C:Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestinal cells (HDL) in plasma.  
 C:Genetics: GDB:APOA1  
 A:Gene: GDB:APOA1  
 A:Cross-references: GDB:119684; OMIM:107680  
 A:Map position: 11q23.3-11q23.3  
 A:Introns: 15/1, 67/2  
 C:Function:  
 A:Description: participates in the reverse transport of cholesterol from tissues to the liver and noncovalently binds and stabilizes prostacyclin (PGI-2)

C:Superfamily: apolipoprotein A-I  
C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipid  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-267/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 100.0%; Score 1362; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-69;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKAAYLTAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
|||||  
Db 1 MKAAYLTAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
OY 61 ALGKQNLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
|||||  
Db 61 ALGKQNLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
OY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHELQEKLSPLGEMRDRARAHV 180  
|||||  
Db 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHELQEKLSPLGEMRDRARAHV 180  
OY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSLSEKAKPALEDLRQ 240  
|||||  
Db 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSLSEKAKPALEDLRQ 240  
OY 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267  
|||||  
Db 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267

## RESULT 2

JS0079  
apolipoprotein A-I precursor - baboon  
C:Species: Papio sp. (baboon)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 13-Jun-1997  
C:Accession: JS0079  
R:Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; Vandenberg, J.L.  
Gene 74, 483-490, 1988  
A:Title: The baboon gene for apolipoprotein A-I: characterization of a cDNA clone and id  
A:Reference number: JS0079; MUID:89232739  
A:Accession: JS0079  
A:Molecule type: mRNA  
A:Residues: 1-267 <HIX>  
A:Experimental source: liver  
C:Comment: This protein is the principal protein component of high density lipoprotein F  
C:Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase co  
C:Comment: This protein contains a region of repeated amino acids which form amphipathic  
C:Genetics:  
A:Gene: apoA1  
C:Superfamily: apolipoprotein A-I  
C:Keywords: HDL; lipid binding; lipoprotein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-267/Product: apolipoprotein A-I #status predicted <LAI>  
F:123-144,145-166,167-188,189-210,211-232,233-254/Region: tandem repeats

Query Match 95.9%; Score 1306; DB 2; Length 267;  
Best Local Similarity 95.1%; Pred. No. 2e-66;  
Matches 254; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 MKAAYLTAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
|||||  
Db 1 MKATVLTAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
OY 61 ALGKQNLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
|||||  
Db 61 ALGKQNLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
OY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHELQEKLSPLGEMRDRARAHV 180  
|||||  
Db 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHELQEKLSPLGEMRDRARAHV 180

OY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSLSEKAKPALEDLRQ 240  
|||||  
Db 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSLSEKAKPALEDLRQ 240  
OY 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267  
|||||  
Db 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267

## RESULT 3

A26529  
apolipoprotein A-I precursor - crab-eating macaque  
C:Species: Macaca fascicularis (crab-eating macaque)  
C:Date: 30-Sep-1989 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
C:Accession: A26529; A26627; S23135; A57766  
R:Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marotli, K.R.  
Gene 49, 103-110, 1986  
A:Title: The primary structure of cynomolgus monkey apolipoprotein A-I deduced from t  
A:Reference number: A26529; MUID:87191989  
A:Accession: A26529  
A:Molecule type: mRNA  
A:Residues: 1-267 <POL>  
A:Cross-references: GB:M15411; NID:g342074; PIDN:AAA36834.1; PID:g342075  
R:Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolo  
Biochemistry 26, 1457-1463, 1987  
A:Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (  
A:Reference number: A26627; MUID:87185451  
A:Accession: A26627  
A:Molecule type: protein  
A:Residues: 25-48 <HER>  
R:Murray, R.W.; Marotli, K.R.  
Biochim. Biophys. Acta 1131, 207-210, 1992  
A:Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and cor  
A:Reference number: S23135; MUID:92305062  
A:Accession: S23135  
A:Molecule type: DNA  
A:Residues: 1-12, 'L', 14-267 <MUR>  
A:Cross-references: GB:M83242; NID:g342070; PIDN:AAA36832.1; PID:g342071  
R:Sorci-Thomas, M.; Kearns, M.W.  
J. Biol. Chem. 266, 18045-18050, 1991  
A:Title: Transcriptional regulation of the apolipoprotein A-I gene.  
A:Reference number: A57766; MUID:92011532  
A:Accession: A57766  
A:Molecule type: DNA  
A:Residues: 1-10 <RES>  
A:Cross-references: GB:M69223; NID:g342066; PIDN:AAA36831.1; PID:g553820  
C:Comment: The precursor is synthesized in the liver and small intestine. The propept  
C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins  
C:Genetics:  
A:Introns: 15/1; 67/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PP1>  
F:25-267/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 95.4%; Score 1299; DB 1; Length 267;  
Best Local Similarity 94.8%; Pred. No. 5e-66;  
Matches 253; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 MKAAYLTAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
|||||  
Db 1 MKATVLTAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
OY 61 ALGKQNLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
|||||  
Db 61 ALGKQNLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
OY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHELQEKLSPLGEMRDRARAHV 180  
|||||



Db 121 VQPYLDDEQKKWQEEEMELYRQKVEPLRAELHEGTRQKLHEHLEKLSPLGEEVDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHEYHAKATEHLSLSEKAKPALLEDLRQ 240  
Db 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHEYHAKASEHLSLSEKAKPALLEDLRQ 240  
QY 241 GLLPVLESFVKVSFLSALEEYTKKLNTQ 267  
Db 241 GLLPVLESFVKVSFLSALEEYTKKLNTQ 267

## RESULT 4

## LDPGAI

apolipoprotein A-I precursor - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 17-Dec-1982 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999  
C:Accession: A60940; A03092; A61418  
R:Luco, C.C.; Li, W.H.; Chan, L.  
J. Lipid Res. 30, 1735-1746, 1989  
A:Title: Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implicati  
A:Reference number: A60940; MUID:90132271  
A:Accession: A60940  
A:Molecule type: mRNA  
A:Residues: 1-266 <LUO>  
R:Chung, H.; Randolph, A.; Reardon, I.; Heintrikson, R.L.  
J. Biol. Chem. 257, 2961-2967, 1982  
A:Title: The covalent structure of apolipoprotein A-I from canine high density lipoprote  
A:Reference number: A03092; MUID:82142425  
A:Accession: A03092  
A:Molecule type: protein  
A:Residues: 25-167, 'G', 169-201, 'Q', 203-234, 'Q', 236-266 <CHU>  
R:Nakai, T.; Whayne, T.F.; Tang, J.  
FEBS Lett. 64, 409-411, 1976  
A:Title: The amino- and carboxyl-terminal sequences of canine apolipoprotein A-I.  
A:Reference number: A61418; MUID:76210910  
A:Accession: A61418  
A:Molecule type: protein  
A:Residues: 25-56, 'Z', 261-262, 'A' <NAK>  
C:Superfamily: apolipoprotein A-I  
C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid transport; li  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 85.3%; Score 1161.5; DB 1; Length 266;  
Best Local Similarity 85.0%; Pred. No. 2.4e-58;  
Matches 227; Conservative 19; Mismatches 20; Indels 1; Gaps 1;  
QY 1 MKAAYLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60  
Db 1 MKAALLTLAVLFLTGSQARHFWQDE-PQSPWDRVKDLATVYVDVAVKDSGRDYVAQFEAS 59  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
Db 60 ALGKQLNLKLDNWDVSTSTFSKLRQIGPTQEFWDNLEKETEVLRQEMSKDLEEVKQK 119  
QY 121 VQPYLDDEQKKWQEEEMELYRQKVEPLRAELQEGARQKLEHLEKLSPLGEEVDRARAHV 180  
Db 120 VQPYLDDEQKKWQEEEMELYRQKVEPLRAELQEGARQKLEHLEKLSPLGEEVDRARAHV 179  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHEYHAKATEHLSLSEKAKPALLEDLRQ 240  
Db 180 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHEYHARASEQLSALGEEKARPALEDLRQ 239  
QY 241 GLLPVLESFVKVSFLSALEEYTKKLNTQ 267  
Db 240 GLLPVLESFVKVSFLSALEEYTKKLNTQ 266

RESULT 5  
LPRB12  
apolipoprotein A-I precursor (clone 2zap AI) - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S06064  
R:Paraskevopoulos, T.B.; Kritsis, A.; Zannis, V.  
submitted to the EMBL data library, July 1989  
A:Reference number: S06064  
A:Accession: S06064  
A:Molecule type: mRNA  
A:Residues: 1-266 <PAR>  
A:Cross-references: EMBL:X15908; NID:g1457; PIDN:CAA34024.1; PID:g1458  
C:Comment: This protein is synthesized in the small intestine.  
C:Comment: This protein is a major component of the high density lipoproteins in plas  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 80.8%; Score 1100.5; DB 1; Length 266;  
Best Local Similarity 80.1%; Pred. No. 6.3e-55;  
Matches 214; Conservative 25; Mismatches 27; Indels 1; Gaps 1;  
QY 1 MKAAYLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60  
Db 1 MKAAYLTLAVLFLTGSQARHFWQDE-PRSSWDKIKDFATVYVDVTKDSGREYVAQFEAS 59  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
Db 60 AFGKQLNLKLDNWDVSTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVROK 119  
QY 121 VQPYLDDEQKKWQEEEMELYRQKVEPLRAELQEGARQKLEHLEKLSPLGEEVDRARAHV 180  
Db 120 VQPYLDDEQKKWQEEEMELYRQKVEPLRAELRESARQKLEHLEKLSPLGEEVDRARAHV 179  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHEYHAKATEHLSLSEKAKPALLEDLRQ 240  
Db 180 DTLRTHLAPYSDELRLQRLAARLEALKENGARLAHEYHAKATEHLSLSEKAKPALLEDLRQ 239  
QY 241 GLLPVLESFVKVSFLSALEEYTKKLNTQ 267  
Db 240 GLLPVLESFVKVSQVNLDEATKKLNTQ 266

RESULT 6  
A46018  
apolipoprotein AI - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
C:Accession: A46018  
R:Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.  
Genomics 15, 643-652, 1993  
A:Title: Characterization of the apolipoprotein AI and CIII genes in the domestic pig  
A:Reference number: A46018; MUID:93224154  
A:Accession: A46018  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-265 <BIR>  
A:Cross-references: GB:L00626; NID:g164358; PIDN:AAA30992.1; PID:g164359  
A:Note: sequence extracted from NCBI backbone (NCBI:129509, NCBI:129511)  
C:Superfamily: apolipoprotein A-I

Query Match 80.8%; Score 1100; DB 2; Length 265;  
Best Local Similarity 80.9%; Pred. No. 6.7e-55;  
Matches 216; Conservative 19; Mismatches 30; Indels 2; Gaps 2;  
QY 1 MKAAYLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60  
Db 1 MKAAYLTLAVLFLTGSQARHFWQDD-PQSPWDRVKDFATVYVDVTKDSGRDYVAQFEAS 59  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
Db 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120









A:Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and A-IV  
A:Reference number: A92558; MUID:87008540  
A:Accession: A24700  
A:Molecule type: DNA  
A:Residues: 1-259 <HAD>  
A:Cross-references: EMBL:J02597; NID:g202935; PIDN:AAA40745.1; PID:g202939  
R:Poncin, J.E.; Martial, J.A.; Gieles, J.E.  
Eur. J. Biochem. 140, 493-498, 1984  
A:Title: Cloning and structure analysis of the rat apolipoprotein A-I cDNA.  
A:Reference number: S00298; MUID:84207987  
A:Accession: S00298  
A:Molecule type: mRNA  
A:Residues: 1-259 <PON>  
A:Cross-references: GB:M00001; EMBL:X00558; NID:g202944; PIDN:AAA40749.1; PID:g202945  
R:Gordon, J.I.; Smith, D.P.; Andy, R.; Alpers, D.H.; Schonfeld, G.; Strauss, A.W.  
J. Biol. Chem. 257, 971-978, 1982  
A:Title: The primary translation product of rat intestinal apolipoprotein A-I mRNA is an  
A:Reference number: A05314; MUID:82098162  
A:Accession: A05314  
A:Molecule type: protein  
A:Residues: 1-18, 'X', 20-21, 'X', 23-28, 'X', 30-38, 'X', 40-41, 'X', 43-45 <GOR>  
C:Comment: This protein is synthesized in the liver and small intestine. The propeptide  
C:Comment: This protein is a major component of the high density lipoproteins in plasma.  
C:Genetics:  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lipoprotein;  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-259/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 62.4%; Score 849.5; DB 2; Length 259;  
Best Local Similarity 63.0%; Pred. No. 6.7e-41;  
Matches 167; Conservative 35; Mismatches 56; Indels 7; Gaps 3;  
QY 1 MKAAYLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVVDVLKDSGRDYVSQFEGS 60  
DB 1 MKAAYLTLAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVVDVLKDSGRDYVSQFEGS 60  
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQQLGPTQEFWDLNLEKETGLRQEMSKDLEEVKAK 120  
DB 60 TLGKQLNLKLLDNWDTLGSTVGRQQLGPTQEFWDLNLEKETGLRQEMSKDLEEVKAK 119  
QY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHDLQEKLSPLGEMRDRARAHV 180  
DB 120 MQPHLDEFQEKWNEVEAYRQKLEPLGTELHKNKAK---EMQRHLKVVAEEFRDMRVNA 175  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAIEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 176 DALRAKFGLYSDQMRNLAQRLTEIKNH--PTLIEYHTRKASDHLKTLGKAKAPALDDLQ 233  
QY 241 GLLPVLESFKVSFLSALLEEYTKLN 265  
DB 234 GLMPVLEAWKAKIMSMIDEAKKLN 258

Search completed: September 22, 2002, 12:06:28  
Job time: 323 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:05:20 ; Search time 35.02 Seconds  
(without alignments)  
295.206 Million cell updates/sec

Title: US-09-803-918A-2  
Perfect score: 1362  
Sequence: 1 MKAAVLTAVLFLTGSQARH.....SFKVSFLSALFEYTKKLTNTQ 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	267	1	APAI_HUMAN
2	1306	95.9	267	1	APAI_MACFA
3	1161.5	85.3	266	1	APAI_CANFA
4	1105	81.1	265	1	APAI_PIG
5	1100.5	80.8	266	1	APAI_RABIT
6	1077	79.1	265	1	APAI_BOVIN
7	975.5	71.6	265	1	APAI_TUPGB
8	904.5	66.4	264	1	APAI_MOUSE
9	842.5	61.9	259	1	APAI_RAT
10	705.5	51.8	264	1	APAI_ANAPL
11	688.5	50.6	264	1	APAI_CHICK
12	663.5	48.7	264	1	APAI_COTJA
13	291	21.4	429	1	APAI_MACFA
14	275	20.2	262	1	APAI_BRARE
15	274.5	20.2	396	1	APAI_BRARE
16	269	19.8	382	1	APAI_PIG
17	268	19.7	391	1	APAI_RAT
18	259	19.0	262	1	APAI_ONCMY
19	251	18.4	395	1	APAI_MOUSE
20	249.5	18.3	401	1	APAI_PAPAN
21	247	18.1	262	1	APAI_ONCMY
22	243.5	17.9	262	1	APAI_SALTR
23	230.5	16.9	260	1	APAI_SPAAU
24	220.5	16.2	258	1	APAI_SALSA
25	209	15.3	311	1	APAI_MOUSE
26	198	14.5	312	1	APAI_RAT
27	190.5	14.0	317	1	APAI_PAPAN
28	189.5	13.9	317	1	APAI_MACFA
29	188	13.8	317	1	APAI_PIG
30	184	13.5	298	1	APAI_CANFA
31	178.5	13.1	305	1	APAI_HUMAN
32	174.5	12.8	317	1	APAI_HUMAN

34	160	11.7	316	1	APE_BOVIN	Q03247	bos taurus
35	147	10.8	311	1	APE_RABIT	P18287	oryctolagus
36	146	10.7	512	1	K2C5_XENLA	P16878	xenopus lae
37	143.5	10.5	771	1	CALD_CHICK	P12957	gallus gall
38	136.5	10.0	482	1	K2C8_RAT	Q10758	rattus norv
39	135	9.9	539	1	MYS3_HYDAT	P39922	hydra atten
40	133.5	9.8	1972	1	MYHB_MOUSE	O08638	mus musculu
41	133	9.8	2230	1	GOG4_HUMAN	Q13439	homo sapien
42	132	9.7	2116	1	MYS2_DICDI	P08799	dictyosteli
43	131	9.6	579	1	G160_HUMAN	Q08378	homo sapien
44	129.5	9.5	502	1	K2C8_XENLA	P08776	xenopus lae
45	128.5	9.4	880	1	RA50_PYRAB	Q9uzc8	pyrococcus

## ALIGNMENTS

RESULT	ID	APAI_HUMAN	STANDARD:	PRT;	267	AA.
AC	P02647;					
DT	21-JUL-1986	(Rel. 01, Created)				
DT	21-JUL-1986	(Rel. 01, Last sequence update)				
DT	16-OCT-2001	(Rel. 40, Last annotation update)				
DE	Apolipoprotein A-I precursor (Apo-AI).					
GN	APOA1.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=84221405; PubMed=6328445;					
RA	Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,					
RA	Baralle F.E.;					
RT	"Human apolipoproteins AI, AII, CII and CIIL. cDNA sequences and mRNA					
RT	abundance.";					
RL	Nucleic Acids Res. 12:3917-3932(1984).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=85026665; PubMed=6207999;					
RA	Selthamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;					
RT	"Isolation and DNA sequence of full-length cDNA and of the entire					
RT	gene for human apolipoprotein AI -- discovery of a new genetic					
RT	polymorphism in the apo AI gene.";					
RL	DNA 3:309-317(1984).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=83220822; PubMed=6406984;					
RA	Shoulders C.C., Kornblith A.R., Munro B.S., Baralle F.E.;					
RT	"Gene structure of human apolipoprotein AI.";					
RL	Nucleic Acids Res. 11:2827-2837(1983).					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=83220772; PubMed=6304641;					
RA	Cheung P., Chan L.;					
RT	"Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";					
RL	Nucleic Acids Res. 11:3703-3715(1983).					
RN	[5]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=84119464; PubMed=6198645;					
RA	Law S.W., Brewer H.B. Jr.;					
RT	"Nucleotide sequence and the encoded amino acids of human					
RT	apolipoprotein A-I mRNA.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).					
RN	[6]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=86008382; PubMed=2995392;					
RA	Law S.W., Brewer H.B. Jr.;					
RT	"Tangier disease. The complete mRNA sequence encoding for					
RT	preproapo-A-I.";					
RL	J. Biol. Chem. 260:12810-12814(1985).					
RN	[7]					



RP SEQUENCE FROM N.A.  
RX MEDLINE=84016011; PubMed=6413973;  
RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
RT "Isolation and characterization of the human apolipoprotein A-I  
gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89377481; PubMed=2673706;  
RA Mogullevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,  
RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
RA Holmquist L., Carlson L.A., Bollen A.;  
RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
coli: purification and biochemical characterization.";  
RL DNA 8:429-436(1989).  
RN [9]  
RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
RX MEDLINE=88196137; PubMed=3129297;  
RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
RA Zannis V.I.;  
RT "Sequence and expression of Tangier apoA-I gene.";  
RL Eur. J. Biochem. 173:465-471(1988).  
RN [10]  
RP SEQUENCE OF 118-267 FROM N.A.  
RX MEDLINE=83091059; PubMed=6294659;  
RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
RT "Isolation and characterization of cDNA clones for human  
apolipoprotein A-I.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
RN [11]  
RP SEQUENCE OF 19-27.  
RX MEDLINE=83256553; PubMed=6409108;  
RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
RA Light J.A.;  
RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
RN [12]  
RP SEQUENCE OF 25-267.  
RX MEDLINE=78123731; PubMed=204308;  
RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
RA Bronzert T.J.;  
RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated  
from high density lipoproteins.";  
RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
RN [13]  
RP SEQUENCE OF 25-267.  
RX MEDLINE=75133493; PubMed=164450;  
RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;  
RT "The primary structure of human plasma high density apolipoprotein  
glutamine I (ApoA-I). II. The amino acid sequence and alignment of  
cyanogen bromide fragments IV, III, and I.";  
RL J. Biol. Chem. 250:2725-2738(1975).  
RN [14]  
RP SEQUENCE OF 25-56.  
RX MEDLINE=88331387; PubMed=3047170;  
RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
A-I (Apo A-I). A novel function of Apo A-I.";  
RL J. Clin. Invest. 82:803-807(1988).  
RN [15]  
RP SEQUENCE OF 25-48.  
RX MEDLINE=89380318; PubMed=2506184;  
RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
RA Chapdelaine A.;  
RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
proteins.";  
RL J. Biol. Chem. 264:16853-16857(1989).  
RN [16]  
RP SEQUENCE OF 25-43.  
RX MEDLINE=88070603; PubMed=3120314;  
RA Prioli R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
RA Pereira M.E.A.;

RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
neuraminidase, to high-density lipoprotein.";  
RL Science 238:1417-1419(1987).  
RN [17]  
RP SEQUENCE OF 25-42.  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [18]  
RP PALMITOYLATION.  
RX MEDLINE=86140194; PubMed=3005308;  
RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
RT "Human apolipoprotein A-I. Post-translational modification by fatty  
acid acylation.";  
RL J. Biol. Chem. 261:3911-3914(1986).  
RN [19]  
RP PROCESSING.  
RX MEDLINE=83195100; PubMed=6405383;  
RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
RA Breslow J.L.;  
RT "Intracellular and extracellular processing of human apolipoprotein  
A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
RN [20]  
RP STRUCTURE BY NMR OF 190-209.  
RX MEDLINE=96270776; PubMed=8664326;  
RA Wang G., Treleaven W.D., Cushley R.J.;  
RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
and CD. Evidence for specific peptide-peptide interactions.";  
RL Biochim. Biophys. Acta 1301:174-184(1996).  
RN [21]  
RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
RX MEDLINE=98024124; PubMed=9356442;  
RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
lipid-bound conformation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
RN [22]  
RP VARIANT MILANO.  
RX MEDLINE=83109095; PubMed=6401735;  
RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,  
RA Franceschini G., Sirtori C.R.;  
RT "Apolipoprotein A-IMilano. Detection of normal A-I in affected  
subjects and evidence for a cysteine for arginine substitution in the  
variant A-I.";  
RL J. Biol. Chem. 258:2508-2513(1983).  
RN [23]  
RP VARIANT TANGIER.  
RX MEDLINE=83300108; PubMed=6412234;  
RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;  
RT "Tangier disease: defective recombination of a specific Tangier  
apolipoprotein A-I isoform (pro-apo A-I) with high density  
lipoproteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).  
RN [24]  
RP VARIANT NORWAY.  
RX MEDLINE=84289383; PubMed=6432779;  
RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,  
RA Utermann G., Haas J., Steinmetz A., Menzel H.J., Assmann G.;  
RT "Abnormal lecithin:cholesterol acyltransferase activation by a human  
apolipoprotein A-I variant in which a single lysine residue is  
deleted.";  
RL J. Biol. Chem. 259:10063-10070(1984).  
RN [25]  
RP SEQUENCE OF 25-107 (VARIANT IOWA).  
RX MEDLINE=89050104; PubMed=3142462;  
RA Nichols W.C., Dwulet F.E., Ilepnieks J., Benson M.D.;  
RT "Variant apolipoprotein AI as a major constituent of a human  
hereditary amyloid.";



Query Match 100.0%; Score 1362; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.8e-71;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYTLAVLFLTGSGARHFWQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
DB 1 MKAAYTLAVLFLTGSGARHFWQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDNLKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDNLKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEMDRARAHV 180  
DB 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEMDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALFEYTKLNTQ 267  
DB 241 GLLPVLESFKVSFLSALFEYTKLNTQ 267

RESULT 2  
APAL\_MACFA STANDARD; PRT; 267 AA.  
AC P15568; P17929;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and  
OS Papio hamadryas (Hamadryas baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541, 9557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.fascicularis;  
RX MEDLINE=87191989; PubMed=3106152;  
RA Polites H.G., Melchior G.W., Castle C.K., Marotti K.R.;  
RT "The primary structure of cynomolgus monkey apolipoprotein A-1  
RL deduced from the cDNA sequence: comparison to the human sequence.";  
RN Gene 49:103-110(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.fascicularis;  
RX MEDLINE=92305062; PubMed=1610902;  
RA Murray R.W., Marotti K.R.;  
RT "Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene  
RL and corresponding flanking regions.";  
RN Biochim. Biophys. Acta 1131:207-210(1992).  
RN [3]  
RP SEQUENCE OF 25-48.  
RC SPECIES=M.fascicularis;  
RX MEDLINE=87185451; PubMed=3105581;  
RA Herbert P.N., Bausserman L.L., Lynch K.M., Saritelli A.L.,  
RA Kantor M.A., Nicolosi R.J., Shulman R.S.;  
RT "Homologues of the human C and A apolipoproteins in the Macaca  
RL fascicularis (cynomolgus) monkey.";  
RN Biochemistry 26:1457-1463(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.hamadryas; TISSUE=Liver;  
RX MEDLINE=89232739; PubMed=2907746;  
RA Hixson J.E., Borenstein S., Cox L.A., Rainwater D.L., Vandeberg J.L.;  
RT "The baboon gene for apolipoprotein A-I: characterization of a cDNA  
RT clone and identification of DNA polymorphisms for genetic studies of

RT cholesterol metabolism.";  
RL Gene 74:483-490(1988).  
RN [5]  
RP SEQUENCE OF 1-10 FROM N.A.  
RC SPECIES=M.fascicularis;  
RA Sorci-Thomas M.;  
RL Submitted (OCT-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS.  
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M15411; AAA36834.1; -.  
CC EMBL; M83242; AAA36832.1; -.  
CC EMBL; M35634; AAA35380.1; -.  
CC EMBL; M69223; AAA36831.1; -.  
CC PIR; A26529; A26529.  
CC PIR; A26627; A26627.  
CC PIR; J50079; J50079.  
CC PIR; S23135; S23135.  
CC HSSP; P02647; 1ODR.  
CC InterPro: IPR000074; Apolipoprotein.  
CC Pfam: PF01442; Apolipoprotein; 1.  
CC Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
CC KW SIGNAL  
CC FT PROPEP 19 24  
CC FT CHAIN 25 267  
CC FT DOMAIN 68 267  
CC FT REPEAT 68 89  
CC FT REPEAT 90 111  
CC FT REPEAT 112 122  
CC FT REPEAT 123 144  
CC FT REPEAT 145 166  
CC FT REPEAT 167 188  
CC FT REPEAT 189 210  
CC FT REPEAT 211 232  
CC FT REPEAT 233 243  
CC FT REPEAT 244 267  
CC FT REPEAT 267 267  
CC FT CONFLICT 13 13 L -> P (IN REF. 1).  
CC FT SEQUENCE 267 AA; 30735 MW; 869955C024088E21 CRC64;

Query Match 95.9%; Score 1306; DB 1; Length 267;  
Best Local Similarity 95.1%; Pred. No. 5.7e-68;  
Matches 254; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 MKAAYTLAVLFLTGSGARHFWQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
DB 1 MKAAYTLAVLFLTGSGARHFWQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDNLKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDNLKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEMDRARAHV 180  
DB 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEMDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESEFKVSFLSALEEYTKKLNTO 267  
|||||  
Db 241 GLLPVLESEFKVSFLSALEEYTKKLNTO 267

RESULT 3  
APAL\_CANFA STANDARD; PRT; 266 AA.  
ID APAL\_CANFA  
AC P02648;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=90132271; PubMed=2515239;  
RA Luo C.-C., Li W.-H., Chan L.;  
RT "Structure and expression of dog apolipoprotein A-I, E, and C-I  
RT mRNAs: implications for the evolution and functional constraints of  
RT apolipoprotein structure."  
RL J. Lipid Res. 30:1735-1746(1989).  
RN [2]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=82142425; PubMed=6801039;  
RA Chung H., Randolph A., Reardon I., Heinrichson R.L.;  
RT "The covalent structure of apolipoprotein A-I from canine high  
RT density lipoproteins."  
RL J. Biol. Chem. 257:2961-2967(1982).  
RN [3]  
RP SEQUENCE OF 25-57 AND 262-265.  
RX MEDLINE=76210910; PubMed=179887;  
RA Nakai T., Whayne T.F., Tang J.;  
RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein  
RT A-I."  
RL FEBS Lett. 64:409-411(1976).  
RN [4]  
RP SEQUENCE OF 25-37.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
RT dog heart proteins."  
RL Electrophoresis 18:2795-2802(1997).  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS. SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
DR PIR; A03092; LPDGA1.  
DR HSSP; P02647; 1ODR.  
DR HSC-2DPAGE; P02648; DOG.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 266  
FT DOMAIN 67 266  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 165

FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 231  
FT REPEAT 232 242  
FT REPEAT 243 266  
FT CONFLICT 168 168  
FT CONFLICT 202 202  
FT CONFLICT 235 235  
FT CONFLICT 264 266  
SQ SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;

Query Match 85.3%; Score 1161.5; DB 1; Length 266;  
Best Local Similarity 85.0%; Pred. No. 9.2e-60;  
Matches 227; Conservative 19; Mismatches 20; Indels 1; Gaps 1;

QY 1 MKAAYTLAVLFLTGSQARHFQWQDEPPQSPWDRVKDLATVYVVDYVKDSGRDYVSQFEQS 60  
|||||  
Db 1 MKAALLTLAVLFLTGSQARHFQWQDE-POSPWDRVKDLATVYVVDYVAVKDSGRDYVAQFEAS 59

QY 61 ALGKQLNKLKLDNWDSTVSTFSKLRQDLPVTOEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
|||||  
Db 60 ALGKQLNKLKLDNWDSTVSTVTKLRQDLPVTOEFWDLNLEKETEVLRQEMSKDLEEVKQK 119

QY 121 VQPYLDDFOKKWQEMELYRQKVEPLRAELQEGARQKHLHELQEKISPLGEEMDRARAHV 180  
|||||  
Db 120 VQPYLDDFOKKWQEMELYRQKVEPLRAELQEGARQKHLHELQEKISPLAEELRDRARTHV 179

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
|||||  
Db 180 DALRAQLAPYSDELRLERLARLEALKENGASLAEYHARASEQLSALGEEKARPALEDLRQ 239

QY 241 GLLPVLESEFKVSFLSALEEYTKKLNTO 267  
|||||  
Db 240 GLLPVLESEFKVSFLSALEEYTKKLNTO 266

RESULT 4  
APAL\_PIG STANDARD; PRT; 265 AA.  
ID APAL\_PIG  
AC P18648;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93224154; PubMed=8468059;  
RA Birchauner A., Knipping G., Juritsch B., Aschauer H., Zechner R.;  
RT "Characterization of the apolipoprotein AI and CIII genes in the  
RT domestic pig."  
RL Genomics 15:643-652(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Brain;  
RX MEDLINE=94125128; PubMed=8294940;  
RA Moeckel B., Zinke H., Flach R., Weiss B., Weller-Guettler H.,  
RA Gassen H.;  
RT "Expression of apolipoprotein A-I in porcine brain endothelium in  
RT vitro."  
RL J. Neurochem. 62:788-798(1994).  
RN [3]  
RP SEQUENCE OF 34-265 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90132667; PubMed=2105375;  
RA Weller-Guettler H., Sommerfeldt M., Papandrikopoulou A., Mischek U.,  
RA Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;  
RT "Synthesis of apolipoprotein A-I in pig brain microvascular

RT endothelial cells.";  
RL J. Neurochem. 54:444-450(1990).  
RN [4]  
RP SEQUENCE OF 105-265 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93154581; PubMed=8428656;  
RA Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;  
RT "Sequences and expression of the porcine apolipoprotein A-I and C-III  
mRNAs.";  
RL Gene 123:173-179(1993).  
RN [5]  
RP SEQUENCE OF 25-265.  
RA Hasler-Rapacz J.O., Choudhary R., Chowdhary B.P., Trieu V.N.,  
RJ Jackson K., Gustavsson I., Rapacz J.;  
RL Submitted (Oct-1995) to the SWISS-PROT data bank.  
RN [6]  
RP SEQUENCE OF 25-34.  
RX MEDLINE=76184721; PubMed=178359;  
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;  
RT "Characterization of the plasma lipoproteins and apoproteins of the  
Erythrocebus patas monkey.";  
RL Biochemistry 15:1928-1933(1976).  
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE LECTITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -I- SUBCELLULAR LOCATION: Extracellular.  
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE  
CC LIVER.  
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L00626; AAA30992.1; -  
DR EMBL; X69477; CAA49234.1; -  
DR EMBL; X17057; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X59414; CAA42050.1; -  
DR PIR; J00704; J00704.  
DR PIR; A05311; A05311.  
DR PIR; A46018; A46018.  
DR PIR; S21830; S21830.  
DR PIR; S31394; S31394.  
DR HSSP; P02647; 1ODR.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 265  
FT DOMAIN 67 265  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 142  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 230  
FT REPEAT 231 241  
FT REPEAT 242 265  
FT REPEAT 265 265  
FT CONFLICT 108 108  
FT CONFLICT 143 143  
FT CONFLICT 173 173  
FT CONFLICT 180 180  
FT CONFLICT 185 186  
FT CONFLICT 209 209  
E -> K (IN REF. 3).  
D -> S (IN REF. 4).  
E -> A (IN REF. 2 AND 3).  
HV -> QL (IN REF. 1 AND 5).  
G -> D (IN REF. 2 AND 3).  
APOLIPOPROTEIN A-I.  
10 X APPROXIMATE TANDEM REPEATS.  
1.  
2.  
3 (HALF-LENGTH).  
4.  
5.  
6.  
7.  
8.  
9 (HALF-LENGTH).  
10.  
MISSING (IN REF. 2 AND 3).  
D -> S (IN REF. 4).  
E -> A (IN REF. 2 AND 3).  
HV -> QL (IN REF. 1 AND 5).  
G -> D (IN REF. 2 AND 3).

FT CONFLICT 224 224 A -> G (IN REF. 4).  
SQ SEQUENCE 265 AA; 30325 MW; 2C6E578318ECF69C CRC64;  
Query Match 81.1%; Score 1105; DB 1; Length 265;  
Best Local Similarity 80.9%; Pred. No. 1.5e-56;  
Matches 216; Conservative 20; Mismatches 29; Indels 2; Gaps 2;  
QY 1 MKAAYLTAVLFLTGSQARHFQODEPPQSDRVKDLATVYVDLKDGRDYSQFEGS 60  
Db 1 MKAAYLTAVLFLTGSQARHFQODEPPQSDRVKDLATVYVDLKDGRDYSQFEGS 59  
QY 61 ALGKQLNLKLDNWDVSTSTESKLRQQLGPTQGFWDNLEKETEGLRQEMSKDLEEVKAK 120  
Db 60 ALGKHLNLKLDNWDVSTSTETKVRQQLGPTQGFWDNLEKETEGLRQEMSKDLEEVKAK 119  
QY 121 VQPYLDDFOKKWQEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMDRARAHV 180  
Db 120 VQPYLDDFOKKWQEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMDRARAHV 179  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAIEYHAKATEHLSISEKAKPALEDLRQ 240  
Db 180 EALRQHVAPYSDDLRLQRLAARLEALKENGARLAIEYHAKATEHLSISEKAKPALEDLRQ 238  
QY 241 GLLPVLESFVSFLSALAEYTKKLNQ 267  
Db 239 GLLPVLENLKVSTLAIDEASKKLNAQ 265  
RESULT 5  
APAL\_RABIT STANDARD; PRT; 266 AA.  
AC P09809;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2ZAP AI; TISSUE=Small intestine;  
RA Paraskevopoulou T.B., Kritis A., Zannis V.I.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=88082866; PubMed=3121329;  
RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,  
RA Kroon P.A., Chao Y.S.;  
RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit  
RT apolipoprotein A-I is synthesized in the intestine but not in the  
RT liver.";  
RL Eur. J. Biochem. 170:99-104(1987).  
RN [3]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=87030294; PubMed=3095115;  
RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;  
RT "The primary structure of apolipoprotein A-I from rabbit high-density  
RT lipoprotein.";  
RL Eur. J. Biochem. 160:427-431(1986).  
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE LECTITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -I- SUBCELLULAR LOCATION: Extracellular.  
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS.  
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----





FT REPEAT 166 187 6.  
FT REPEAT 188 209 7.  
FT REPEAT 210 230 8.  
FT REPEAT 231 241 9 (HALF-LENGTH).  
FT REPEAT 242 265 10.  
FT CONFLICT 185 186 OL -> HV (IN REF. 2).  
SQ SEQUENCE 265 AA; 30276 MW; 06A2681EA2ABA50F CRC64;

Query Match 79.1%; Score 1077; DB 1; Length 265;  
Best Local Similarity 78.7%; Pred. No. 5.7e-55;  
Matches 210; Conservative 23; Mismatches 32; Indels 2; Gaps 2;

QY 1 MKAAVLTAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
DB 1 MKAAVLTAVLFLTGSQARHFWQDD-PQSSWDRVKDFATVYVEAIKDSGRDYVAQFEAS 59  
QY 61 ALGKQNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
DB 60 ALGKQNLKLLDNWDTLASTLSKVRQGLPVTQEFWDLNLEKETASLRQEMHKDLEEVKQK 119  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHEIQEKLSPLEEMDRARAHV 180  
DB 120 VQPYLDDFQKKWHEVEIYRQKVAPLGEEFREGARQKVELQDKLSPLAQELRDRARAHV 179  
QY 181 DALRTHLAPYSDELRLRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 180 ETLRQDLAPYSDDLRLRLARLEALKENGGS-LAEYHAKASEQLKALGEKAKPYLEDLRQ 238  
QY 241 GLLPVLESFKVSFLSALAEYTKKINTQ 267  
DB 239 GLLPVLESFKVSFLSALAEYTKKINTQ 265

RESULT 7  
APAL\_TUPGB STANDARD; PRT; 265 AA.  
ID APAL\_TUPGB  
AC 018759;

DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Tupaiia glis belangeri (Common tree shrew).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.  
OX NCBI\_TaxID=9396;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBD databases.

CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE EFFLUX OF CHOLESTEROL AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF005638; AAB82326.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
CC Pfam; PF01442; Apolipoprotein; 1.

KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 265 BY SIMILARITY.  
FT DOMAIN 67 265 APOLOPOPROTEIN A-I.  
FT REPEAT 67 88 10 X APPROXIMATE TANDEM REPEATS.  
FT REPEAT 89 110 1.  
FT REPEAT 111 121 2.  
FT REPEAT 122 143 3 (HALF-LENGTH).  
FT REPEAT 144 165 4.  
FT REPEAT 166 187 5.  
FT REPEAT 188 209 6.  
FT REPEAT 210 231 7.  
FT REPEAT 232 242 8.  
FT REPEAT 243 265 9 (HALF-LENGTH).  
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAEA165 CRC64;

Query Match 71.6%; Score 975.5; DB 1; Length 265;  
Best Local Similarity 69.7%; Pred. No. 3.3e-49;  
Matches 184; Conservative 38; Mismatches 41; Indels 1; Gaps 1;

QY 1 MKAAVLTAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
DB 1 MKAAVLTAVLFLTGSQARHFWQDE-PQSSWDRVKDLATVYVDVAVKESGREYVSQLEAS 59  
QY 61 ALGKQNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
DB 60 ALGKQNLKLLDNWDTLGSTFQKVHHLGPAVQEFWEKLEKETELRREINKDLEDVRQK 119  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHEIQEKLSPLEEMDRARAHV 180  
DB 120 TQPYLDDFQKKWQEDDLERYRQKVEPLRAELQEGARQKLHEIQEKLSPLEEMDRARAHV 179  
QY 181 DALRTHLAPYSDELRLRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 180 DTLRTQLAPYSEDMRKTLGARLEAIKEGSSASLAEYHAKASEQLSALGEKAKPYLEDLRQ 239  
QY 241 GLLPVLESFKVSFLSALAEYTKK 264  
DB 240 GLMPWESFKVGLNVIDEAAKTL 263

RESULT 8  
APAL\_MOUSE STANDARD; PRT; 264 AA.  
ID APAL\_MOUSE  
AC 000623;

DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE EFFLUX OF CHOLESTEROL AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF005638; AAB82326.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
CC Pfam; PF01442; Apolipoprotein; 1.



```
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECTITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC EMBL; X64262; CAA45560.1; -.
CC EMBL; X64263; CAA45561.1; -.
CC EMBL; L04149; -; NOT_ANNOTATED_CDS.
CC EMBL; L04151; -; NOT_ANNOTATED_CDS.
CC PIR; S22420; S22420.
CC PIR; A44364; A44364.
CC HSSP; P02647; LAV1.
CC SWISS-2DPAGE; Q00623; MOUSE.
CC MGD; MGT:88049; ApoA1.
CC InterPro; IPR000074; Apolipoprotein.
CC pfam; PF01442; Apolipoprotein; 1.
CC Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; signal.
CC SIGNAL 1 18 BY SIMILARITY.
CC PROPEP 19 24 BY SIMILARITY.
CC CHAIN 25 264 APOLIPOPROTEIN A-I.
CC DOMAIN 67 264 10 X APPROXIMATE TANDEM REPEATS.
CC REPEAT 67 88 1.
CC REPEAT 89 110 2.
CC REPEAT 111 121 3 (HALF-LENGTH).
CC REPEAT 122 143 4.
CC REPEAT 144 165 5.
CC REPEAT 166 187 6.
CC REPEAT 188 207 7 (INCOMPLETE).
CC REPEAT 208 229 8.
CC REPEAT 230 240 9 (HALF-LENGTH).
CC REPEAT 241 264 10.
CC SEQUENCE 264 AA; 30587 MW; C453FF2019634AAC CRC64;

Query Match 66.4%; Score 904.5; DB 1; Length 264;
Best Local Similarity 65.2%; Pred. No. 3.6e-45;
Matches 174; Conservative 41; Mismatches 49; Indels 3; Gaps 2;

QY 1 MKAAVLTAVFLTGSQARHFWQDEPQSPWDRVKDLATVYVDLKDGRDYSQEGS 60
   ||| | :||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MKAVVLAVAVFLTGSQAWHWQODE-PQSQWDKVKDFANVYVDVAVKDSGRDYSQFESS 59
QY 61 ALGKOLNLKLDNWDVSTSTFSKLRQIGPVTOEFWDNLKETEGLRQEMSKDLEEVAK 120
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 60 SLGQOLNLNLEWMDTLGTVSOLQERIGPLTRDFWDLNLEKETDWWKQEMNKDLEEVAK 119
QY 121 VQPYLDDFOKKWQEMELYRQKVEPLRAFLQEGARQKLHELQKLSPLGEEMDRARAHV 180
   ||||| :||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 VQPYLDEFQKKMREDELRYRQKVAPLGAELQESARQKLOELQGRLSVAEEFRDRMRTHV 179
QY 181 DALRTHLAPYSDELRLARLARLEALKENGARLARLAHYHAKATEHTLSEKAKPALEDLRQ 240
   | :||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 DSLRTQLAPHSEQMRSLAQRLAELKSN--PLTNEYHTRAKTHLKLGEKARPALEDLRH 237
QY 241 GLLPVLESEKVSFLSALBEYTKLINTQ 267
   | :||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 238 SIMPMLETLTKRAQSVIDKASETLTAQ 264

RESULT 9
APAL_RAT
ID APAL_RAT STANDARD; PRT; 259 AA.
AC P04639;
```

```
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84207987; Pubmed=6426956;
RA Poncin J.E., Martial J.A., Gjelten J.E.;
RT "Cloning and structure analysis of the rat apolipoprotein A-I cDNA.";
RL Eur. J. Biochem. 140:493-498(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87008540; Pubmed=3020028;
RA Haddad I.A., Ordovas J.M., Fitzpatrick T., Karathanasis S.K.;
RT "Linkage, evolution, and expression of the rat apolipoprotein A-I, C-
RT III, and A-IV genes.";
RL J. Biol. Chem. 261:13268-13277(1986).
RN [3]
RP SEQUENCE OF 1-45.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=82098162; Pubmed=6798036;
RA Gordon J.I., Smith D.P., Alpers D.H., Schonfeld G.,
RA Strauss A.W.;
RT "The primary translation product of rat intestinal apolipoprotein A-I
RT mRNA is an unusual preproprotein.";
RL J. Biol. Chem. 257:971-978(1982).
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL EFFLUX FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC THE LECTITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC EMBL; M00001; AAA40749.1; -.
CC EMBL; X00558; CAA25224.1; -.
CC EMBL; J02597; AAA40745.1; -.
CC PIR; A24700; A24700.
CC HSSP; P02647; LAV1.
CC InterPro; IPR000074; Apolipoprotein.
CC pfam; PF01442; Apolipoprotein; 1.
CC Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; signal.
CC SIGNAL 1 18
CC PROPEP 19 24 APOLIPOPROTEIN A-I.
CC CHAIN 25 259 10 X APPROXIMATE TANDEM REPEATS.
CC DOMAIN 67 259 1.
CC REPEAT 67 88 1.
CC REPEAT 89 110 2.
CC REPEAT 111 121 3 (HALF-LENGTH).
CC REPEAT 122 143 4.
CC REPEAT 144 161 5 (INCOMPLETE).
CC REPEAT 162 183 6.
CC REPEAT 184 203 7 (INCOMPLETE).
CC REPEAT 204 225 8.
CC REPEAT 226 236 9 (HALF-LENGTH).
CC REPEAT 237 259 10.
CC REPEAT 201 201 R -> K (IN REF. 2).
CC CONFLICT 214 214 G -> S (IN REF. 2).
CC CONFLICT 218 218 R -> K (IN REF. 2).
```



RP SEQUENCE FROM N.A.  
RX MEDLINE=92381402; PubMed=1512510;  
RA Lamou-Fava S., Sastry R., Ferrari S., Rajavashisth T.B.,  
RA Lusis A.J., Karathanasis S.K.;  
RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene  
RT expression: differences between avian and mammalian apoA-I gene  
RT transcription control regions."  
RL J. Lipid Res. 33:831-842(1992).  
RN [5]  
RP SEQUENCE OF 25-44.  
RX MEDLINE=83213468; PubMed=6406496;  
RA Shackleford J.E., Leberer H.G.;  
RT "Synthesis and secretion of apolipoprotein A1 by chick breast  
RT muscle."  
RL J. Biol. Chem. 258:7175-7180(1983).  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M17961; AAA48593.1; -  
DR EMBL; M18746; AAA48594.1; -  
DR EMBL; M25559; AAA48592.1; -  
DR EMBL; M96012; AAA48597.1; -  
DR PIR; S01453; LPCHAL.  
DR PIR; JH0471; JH0471.  
DR HSSP; P02647; 1ODR.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 264  
FT DOMAIN 67 264  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 231  
FT REPEAT 232 242  
FT REPEAT 243 264  
FT CONFLICT 16 16 T -> I (IN REF. 4).  
FT CONFLICT 148 148 E -> K (IN REF. 3).  
SQ SEQUENCE 264 AA; 30680 MM; 968320E81E2AC5C2 CRC64;  
  
Query Match 50.6%; Score 688.5; DB 1; Length 264;  
Best Local Similarity 50.0%; Pred. No. 6.7e-33;  
Matches 132; Conservative 57; Mismatches 74; Indels 1; Gaps 1;  
  
QY 1 MKAVALTLAVLFLTGSAARHFWQDEPPQSPWDRVKDLATVYVDYDKDSGRDYSQFEES 60  
DB 1 MRGVLVTLAVLFLTGTAARHFWQDEPPQSPWDRVKDLATVYVDYDKDSGRDYSQFEES 59  
QY 6T ALGKQLNKLKLDNWDVSTFSTKRLREGDGPVTOEFWDNLEKETEGLRQEMSKDLEEVAK 120  
DB 60 AVGKQLDLKLDNWDVSTFSTKRLREGDGPVTOEFWDNLEKETEGLRQEMSKDLEEVAK 119

QY 121 VQPLDDFQKKQWQEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMDRARAHV 180  
DB 120 IRPFLDQFSAKWTEELQYRORLTPVAQELKELTKQKVELMQAKLTPVAEARDRLRGHV 179  
QY 181 DALRTHLAPYSDELQRRLARLEALKENGARLAEXHAKATEHLSLSEKAPPALEDLRQ 240  
DB 180 EELRNULAPYSDELQRRLARLEALKENGARLAEXHAKATEHLSLSEKAPPALEDLRQ 239  
QY 241 GLPLVLESEFKVSFLSALEETTKL 264  
DB 240 RLTPYAENLKNRLISFLDELQKSV 263  
  
RESULT 12  
APAL\_COTJA STANDARD; PRT; 264 AA.  
ID APAL\_COTJA  
AC P32918;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Coturnix coturnix japonica (Japanese quail).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
OX NCBI\_TaxID=93934;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=97212027; PubMed=9058967;  
RA Oku H., Toda T., Nagata J., Ishikawa M., Neyazaki K., Shinyo C.,  
RA Chinen I.;  
RT "Apolipoprotein A-1 of Japanese quail: cDNA sequence and modulation of  
RT tissue expression by cholesterol feeding."  
RL Biosci. Biotechnol. Biochem. 61:286-290(1997).  
RN [2]  
RP SEQUENCE OF 25-60.  
RX MEDLINE=93213845; PubMed=8461329;  
RA Oku H., Ishikawa M., Nagata J., Toda T., Chinen I.;  
RT "Lipoprotein and apoprotein profile of Japanese quail."  
RL Biochim. Biophys. Acta 1167:22-28(1993).  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF VLDL, HDL, LDL AND IN  
CC CHYLOMICRONS. EXPRESSED IN A NUMBER OF TISSUES INCLUDING LIVER,  
CC SMALL INTESTINE, LUNG, KIDNEY, HEART AND MUSCLE WITH HIGHEST  
CC EXPRESSION IN LIVER AND SMALL INTESTINE.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D85133; BAA12729.1; -  
DR HSSP; P02647; 1GM4.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 264  
FT DOMAIN 67 264  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 121 121 3 (HALF-LENGTH).  
FT REPEAT 121 121 3 (HALF-LENGTH).





```

RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97385153; PubMed=9238027;
RA Babln P.J., Thisse C., Durliat M., Andre M., Akimenko M.-A.,
RA Thisse B.;
RT "Both apolipoprotein E and A-I genes are present in a nonmammalian
RT vertebrate and are highly expressed during embryonic development.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8622-8627(1997).
CC -I- FUNCTION: APOA-I PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT) (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC -I- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN THE YOLK SYNCYTIAL LAYER
CC DURING EMBRYONIC (STARTING AT THE GASTRULA STAGE) AND EARLY LARVAL
CC DEVELOPMENT, AN EXTRAMEMBRYONIC STRUCTURE IMPLICATED IN EMBRYONIC
CC AND LARVAL NUTRITION.
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y13653; CAA74004.1; -.
DR ZFIN; ZDB-GENE-990415-14; apoA.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 23 POTENTIAL.
FT CHAIN 24 262 APOLIPOPROTEIN A-I.
FT DOMAIN 32 63 3 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 64 262 10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 64 85 1.
FT REPEAT 87 107 2.
FT REPEAT 108 118 3 (HALF-LENGTH).
FT REPEAT 119 140 4.
FT REPEAT 141 162 5.
FT REPEAT 163 184 6.
FT REPEAT 185 206 7.
FT REPEAT 207 228 8.
FT REPEAT 229 239 9 (HALF-LENGTH).
FT REPEAT 240 262 10.
SQ SEQUENCE 262 AA; 30256 MW; BB839A0A815365B9 CRC64;

Query Match 20.2%; Score 275; DB 1; Length 262;
Best Local Similarity 25.8%; Pred. No. 2e-09;
Matches 68; Conservative 61; Mismatches 131; Indels 4; Gaps 3;

```

```

RESULT 15
AP04_HUMAN
ID AP04_HUMAN STANDARD; PRT; 396 AA.
AC P06727;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV).
GN APOA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=89194198; PubMed=2930771;
RA Yang C., Gu Z.W., Xiong W., Rosseneu M., Yang H.X., Lee B.M.,
RA Gotto A.M. Jr., Chan L.;
RT "The primary structure of human apolipoprotein A-IV.";
RL Biochim. Biophys. Acta 1002:231-237(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87041474; PubMed=3095836;
RA Karathanasis S.K., Oeltgen P., Haddad I.A., Antonarakis S.E.;
RT "Structure, evolution, and polymorphisms of the human apolipoprotein
RT A4 gene (APOA4).";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86296629; PubMed=3755616;
RA Karathanasis S.K., Yunis I.;
RT "Structure, evolution, and tissue-specific synthesis of human
RT apolipoprotein AIV.";
RL Biochemistry 25:3962-3970(1986).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=87250378; PubMed=3036793;
RA Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,
RA Gordon J.I., Taylor J.M.;
RT "Structure and expression of the human apolipoprotein A-IV gene.";
RL J. Biol. Chem. 262:7973-7981(1987).
RN [5]
RP SEQUENCE OF 21-396 FROM N.A.
RX MEDLINE=86111885; PubMed=3080432;
RA Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;
RT "The nucleotide and derived amino acid sequence of human
RT apolipoprotein A-IV mRNA and the close linkage of its gene to the
RT genes of apolipoproteins A-I and C-III.";
RL J. Biol. Chem. 261:1998-2002(1986).
RN [6]
RP SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=84161950; PubMed=6706947;
RA Gordon J.I., Bisgaier C.L., Sims H.F., Sachdev O.P., Glickman R.M.,
RA Strauss A.W.;
RT "Biosynthesis of human preapolipoprotein A-IV.";
RL J. Biol. Chem. 259:468-474(1984).
RN [7]
RP REVIEW ON POLYMORPHISM.
RA Lohse P., Brewer H.B. Jr.;
RT "Genetic polymorphism of apolipoprotein A-IV.";
RL Curr. Opin. Lipidol. 2:90-95(1991).
RN [8]
RP VARIANT A-IV*2.
RX MEDLINE=90277616; PubMed=2351649;
RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;
RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to
RT nucleotide substitutions in the apolipoprotein A-IV gene.";
RL J. Biol. Chem. 265:10061-10064(1990).
RN [9]
RP VARIANTS A-IV*0 AND A-IV*3.

```



RX MEDLINE=90324273; PubMed=1973689;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for  
 RT two rare variants of apolipoprotein A-IV-1.";  
 RL J. Biol. Chem. 265:12734-12739(1990).  
 RN [10]  
 RP VARIANTS.  
 RX MEDLINE=91310615; PubMed=1677358;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
 RT 1(Thr-347-->Ser), apoA-IV-0(Lys-167-->Glu,Gln-360-->His), and apoA-IV-  
 RT 3(Glu-165-->Lys).";  
 RL J. Biol. Chem. 266:13513-13518(1991).  
 RN [11]  
 RP ERRATUM.  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RL J. Biol. Chem. 266:19866-19866(1991).  
 RN- [12]  
 RP VARIANT MET-13.  
 RX MEDLINE=92238494; PubMed=1349197;  
 RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
 RA Assmann G.;  
 RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
 RT gene are associated with changes in the concentration of apo B- and  
 RT apo A-I-containing lipoproteins in a normal population.";  
 RL Am. J. Hum. Genet. 50:1115-1128(1992).  
 RN [13]  
 RP VARIANT SER-147.  
 RX MEDLINE=92144647; PubMed=1737067;  
 RA Tenkanen H., Koskinen P., Metsä J., Baumann M., Lukka M.,  
 RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
 RA Manninen V., Ehnholm C.;  
 RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
 RT asparagine to serine substitution at residue 127.";  
 RL Biochim. Biophys. Acta 1138:27-33(1992).  
 RN [14]  
 RP VARIANT A-IV\*5.  
 RX MEDLINE=93138374; PubMed=1487136;  
 RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
 RA Ferrell R.E., Pollitzer W.S.;  
 RT "Molecular basis of a unique African variant (A-IV 5) of human  
 RT apolipoprotein A-IV and its significance in lipid metabolism.";  
 RL Genet. Epidemiol. 9:379-388(1992).  
 RN [15]  
 RP VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
 RX MEDLINE=95245341; PubMed=7728150;  
 RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
 RA Csaszar A.;  
 RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
 RT frequencies, effect on lipid levels, and sequence of two new  
 RT variants.";  
 RL Hum. Mutat. 5:58-65(1995).  
 RN [16]  
 RP VARIANTS FCHL SEATTLE SER-161; LEU-178 AND GLN-264.  
 RX MEDLINE=97114287; PubMed=8956036;  
 RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;  
 RT "Two novel apolipoprotein A-IV variants in individuals with familial  
 RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
 RT activity.";  
 RL Hum. Mutat. 8:319-325(1996).  
 CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
 CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
 CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
 CC COMPONENT OF HDL AND CHYLOMICRONS.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
 CC SECRETED IN PLASMA.  
 CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH  
 CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
 CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
 CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
 CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL  
 CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.

```
CC -I- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV*0  
CC TO APOA-IV*7). APOA-IV*1 IS THE MAJOR ALLELE (90%), IV*2 IS ALSO  
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.  
CC -I- DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN  
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF  
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).  
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce,  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; M14642; AAA51745.1; -.  
DR EMBL; X13629; CAA31955.1; -.  
DR EMBL; M14566; AAA51748.1; -.  
DR EMBL; J02758; AAAG6731.1; -.  
DR EMBL; M13654; AAA51744.1; -.  
DR PIR; A26481; LPHUO4.  
DR PIR; A24449; A24449.  
DR PIR; A29330; A29330.  
DR PIR; A26280; A26280.  
DR PIR; S02715; S02715.  
DR HSSP; P02649; INFO.  
DR SWISS-2DPAGE; P06727; HUMAN.  
DR MIM; 107690; -.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;  
KW Disease mutation; Polymorphism.  
FT SIGNAL 1 20  
FT CHAIN 21 396 APOLIPOPROTEIN A-IV.  
FT DOMAIN 33 330 13 x 22 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 33 54 1.  
FT REPEAT 60 81 2.  
FT REPEAT 82 103 3.  
FT REPEAT 115 136 4.  
FT REPEAT 137 158 5.  
FT REPEAT 159 180 6.  
FT REPEAT 181 202 7.  
FT REPEAT 203 224 8.  
FT REPEAT 225 246 9.  
FT REPEAT 247 268 10.  
FT REPEAT 269 286 11.  
FT REPEAT 287 308 12.  
FT REPEAT 309 330 13.  
FT DOMAIN 372 389 GLU/GLN-RICH.  
FT VARIANT 13 V->M (IN APOA-IV*1D).  
FT FTId=VAR_000626.  
FT E->K (IN BUDAPEST-2).  
FT VARIANT 44 44 /FTId=VAR_000627.  
FT N->S (IN APOA-IV*1B).  
FT FTId=VAR_000628.  
FT A->S (IN SEATTLE-3; IN FCHL).  
FT VARIANT 161 161 /FTId=VAR_000629.
```

Db 113 LLPHANEVSQKIGDNLRELOQRLEPYADQLRTQVNTQAEQLRRQOLTPYAQRMERVLRENA 172  
QY 181 DALRTHLAPYSDELROQLARLEALKENGARLA EYHAKATEHLSTLSEKAKPALEDLRQ 240  
Db 173 DSLQASLRPHADELKAIDQNVLELK-----GRLTPYADEFKVKIDQTVLEELRR 221  
QY 241 GLLPVLESFKVSFLSALEEY 261  
Db 222 SLAPYAQDTQOEKLNHQLEGLT 242

Search completed: September 22, 2002, 12:22:16  
Job time: 1016 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:03:00 ; Search time 111.17 Seconds  
(without alignments)  
415.487 Million cell updates/sec

Title: US-09-803-918A-2  
Perfect score: 1362  
Sequence: 1 MKAVALTLAVLFLTGSOARH.....SPKVSFLSALLEYTKKLTNQ 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvivirus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeap:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1012.5	74.3	264	11 Q9Z2L4	Q9Z2L4 mesocricetu
2	896	65.8	263	11 Q08855	Q08855 mus musculu
3	895	65.7	263	11 Q09042	Q09042 mus musculu
4	836	61.4	258	11 Q09054	Q09054 rattus norv
5	834	61.2	258	11 Q08877	Q08877 rattus norv
6	727.5	53.4	241	6 Q9TS49	Q9TS49 erinaceus e
7	327	24.0	67	4 Q9Y355	Q9Y355 homo sapien
8	312.5	22.9	82	6 Q29248	Q29248 sus scrofa
9	298.5	21.9	263	13 Q98TG6	Q98TG6 anguilla ja
10	257.5	18.9	366	13 Q93601	Q93601 gallus gall
11	253	18.6	435	11 Q01488	Q01488 mus musculu
12	247	18.1	56	6 Q02762	Q02762 ovis aries
13	247	18.1	395	11 Q91XF8	Q91XF8 mus musculu
14	244	17.9	395	11 Q9DBN0	Q9DBN0 mus musculu
15	213.5	15.7	275	13 Q9PT02	Q9PT02 oncorhynch
16	200	14.7	363	4 Q9UBJ3	Q9UBJ3 homo sapien

17	194.5	14.3	1547	5 Q26471	Q26471 schistocerc
18	191	14.0	313	6 Q9GLC0	Q9GLC0 tupaiia glis
19	185	13.6	244	4 Q13784	Q13784 homo sapien
20	184.5	13.5	317	6 Q9GLM6	Q9GLM6 hylobates l
21	181.5	13.3	317	6 Q9GLM8	Q9GLM8 gorilla gor
22	181	13.3	259	13 Q98TG5	Q98TG5 anguilla ja
23	179.5	13.2	317	6 Q9GJ03	Q9GJ03 pan troglod
24	178	13.1	259	13 Q98TG4	Q98TG4 anguilla ja
25	177.5	13.0	259	13 Q98TG3	Q98TG3 anguilla ja
26	177.5	13.0	317	6 Q9GLM7	Q9GLM7 pongo pygma
27	171	12.6	367	11 Q9QUH3	Q9QUH3 rattus norv
28	171	12.6	368	11 Q99P64	Q99P64 mus musculu
29	170.5	12.5	259	13 Q98TG2	Q98TG2 anguilla ja
30	170	12.5	368	11 Q91X90	Q91X90 mus musculu
31	169	12.4	174	13 Q98S13	Q98S13 cyprinus ca
32	164.5	12.1	259	13 Q98TG1	Q98TG1 anguilla ja
33	153.5	11.3	174	13 Q9DFQ3	Q9DFQ3 gillichthys
34	146	10.7	1478	4 Q9BQS8	Q9BQS8 homo sapien
35	142	10.4	1411	4 Q15075	Q15075 homo sapien
36	141	10.4	26	4 Q9UCT8	Q9UCT8 homo sapien
37	141	10.4	1456	5 Q9V587	Q9V587 drosophila
38	139.5	10.2	194	11 Q9D2P9	Q9D2P9 mus musculu
39	137	10.1	470	5 Q9XTM4	Q9XTM4 caenorhabdi
40	137	10.1	1935	5 Q44934	Q44934 loligo peal
41	136	10.0	992	4 Q9NTH6	Q9NTH6 homo sapien
42	136	10.0	1410	4 Q14221	Q14221 homo sapien
43	136	10.0	1455	4 Q9UPV0	Q9UPV0 homo sapien
44	135.5	9.9	1047	11 Q9EP81	Q9EP81 mus musculu
45	135.5	9.9	1708	5 Q9U0S6	Q9U0S6 mytilus gal

ALIGNMENTS

RESULT	ID	Q9Z2L4	PRELIMINARY;	PRT;	264 AA.
AC	Q9Z2L4	Q9Z2L4			
DT	01-MAY-1999	(TREMBLREL. 10, Created)			
DT	01-MAY-1999	(TREMBLREL. 10, Last sequence update)			
DT	01-JUN-2001	(TREMBLREL. 17, Last annotation update)			
DE	APOLIPROTEIN A-I.				
GN	APOAI.				
OS	Mesocricetus auratus (Golden hamster).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;				
OC	Mesocricetus.				
OX	NCBI_TaxID=10036;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;				
RX	MEDLINE=99061559; PubMed=9843713;				
RA	Wu J.Y.J., Reaves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;				
RT	"Zinc deficiency decreases plasma level and hepatic mRNA abundance of				
RT	apolipoprotein A-I in rats and hamsters."				
RL	Am. J. Physiol. 275:C1516-C1525(1998).				
DR	EMBL; AF046919; AAC98484.1; -				
DR	HSSP; P02647; LAVI.				
DR	InterPro; IPR000074; Apolipoprotein.				
DR	Pfam; PF01442; Apolipoprotein; 1.				
KW	Lipoprotein.				
SQ	SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;				

Query Match 74.3%; Score 1012.5; DB 11; Length 264;  
Best Local Similarity 73.0%; Pred. No. 8.7e-56;  
Matches 195; Conservative 30; Mismatches 39; Indels 3; Gaps 2;

QY	1	MKAVALTLAVLFLTGSOARHFWQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
DB	1	MKTIVLAVAVLFLTGSOARHFWQDD-PQTPMDRVKDFATVYVDVAVKDSGREYVSQFETS 59
QY	61	ALGKQLNLKLLDNWDSVSTFSKLRQLGPTVQEFWNLEKETEGLRQEMSKDLEEVKAK 120

DB 60 ALGKQLNLNLEMDTLGSTVGRLOEQLGVTQEFWDLNLEKETEMLRREMKNKDLLEEVKAK 119  
121 VQPYLDDEFOKKWQEMELLYRQKVEPLRAELQEGAROKLHELOEKLSPLEEMRDRARAHV 180  
120 VQPYLDQFQTKWQEEVALYRQKMEPLGAELRDGAQOKLOEQLEKLTPLGEDLRDRMRHHV 179  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 180 DALRTKMTPTPSDQMDRLAERLAQLKDS--PTLAEYHTKAADHLKAFGEKAKPALEDLRQ 237  
QY 241 GLLPVLESEFKVSFLSALEEYTKLNTQ 267  
DB 238 GLMPVFESFKTRIMSMVEEASKKLNAQ 264  
RESULT 2  
ID 008855 PRELIMINARY; PRT; 263 AA.  
AC 008855;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=SPLEEN;  
RX MEDLINE=98077648; PubMed=9415807;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
gene.";  
RL Biochem. Mol. Biol. Int. 43:989-996(1997).  
DR EMBL; U79574; AAB58426.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 263 AA; 30544 MW; 7825DF7483A5B24A CRC64;

Query Match 65.8%; Score 896; DB 11; Length 263;  
Best Local Similarity 65.2%; Pred. No. 1.5e-48;  
Matches 174; Conservative 41; Mismatches 48; Indels 4; Gaps 3;

QY 1 MKAAYLTLAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQEGS 60  
DB 1 MKAAYLAVLAVLFLTGSQAMHWQODE-PQSQWDKVKDFANVYVDVAVKDSGRDYVSQFESS 59  
QY 61 ALGKQLNLKLDNDWSDVSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
DB 60 SLGQQLNLNLEMDTLGSTVGRLOEQLEGLPLTRDFWDLNLEKETDWRQEMKNKDLLEEVKOK 119  
QY 121 VQPYLDDEFOKKWQEMELLYRQKVEPLRAELQEGAROKLHELOEKLSPLEEMRDRARAHV 180  
DB 120 VQPYLDQFQTKWQEEVALYRQKMEPLGAELRDGAQOKLOEQLEKLTPLGEDLRDRMRHHV 178  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 179 DSLRTQLAPHSEQMRSLAQLRAELKSN--PTLNEYHTRAKTHLKTGKAKPALEDLRH 236  
QY 241 GLLPVLESEFKVSFLSALEEYTKLNTQ 267  
DB 237 SIMPMLFTLKTQVQSVIDKASETLTAQ 263  
RESULT 3  
ID 009042 PRELIMINARY; PRT; 263 AA.  
AC 009042;

DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
DE APOLIPROTEIN A-I.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ICR AND BALB/C, AND C3H; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79575; AAB58427.1; -.  
DR EMBL; U79572; AAB58424.1; -.  
DR EMBL; U79573; AAB58425.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 263 AA; 30516 MW; E245DF7483A5A0DD CRC64;

Query Match 65.7%; Score 895; DB 11; Length 263;  
Best Local Similarity 65.2%; Pred. No. 1.8e-48;  
Matches 174; Conservative 41; Mismatches 48; Indels 4; Gaps 3;

QY 1 MKAAYLTLAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQEGS 60  
DB 1 MKAAYLAVLAVLFLTGSQAMHWQODE-PQSQWDKVKDFANVYVDVAVKDSGRDYVSQFESS 59  
QY 61 ALGKQLNLKLDNDWSDVSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
DB 60 SLGQQLNLNLEMDTLGSTVGRLOEQLEGLPLTRDFWDLNLEKETDWRQEMKNKDLLEEVKOK 119  
QY 121 VQPYLDDEFOKKWQEMELLYRQKVEPLRAELQEGAROKLHELOEKLSPLEEMRDRARAHV 180  
DB 120 VQPYLDQFQTKWQEEVALYRQKMEPLGAELRDGAQOKLOEQLEKLTPLGEDLRDRMRHHV 178  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 179 DSLRTQLAPHSEQMRSLAQLRAELKSN--PTLNEYHTRAKTHLKTGKAKPALEDLRH 236  
QY 241 GLLPVLESEFKVSFLSALEEYTKLNTQ 267  
DB 237 SIMPMLFTLKTQVQSVIDKASETLTAQ 263

RESULT 4  
ID 009054 PRELIMINARY; PRT; 258 AA.  
AC 009054;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
DE APOLIPROTEIN A-I.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WKY, AND SHRSP; TISSUE=SPLEEN;  
RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79578; AAB58430.1; -.  
DR EMBL; U79577; AAB58429.1; -.  
DR HSSP; P02647; IAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW lipoprotein.  
SQ SEQUENCE 258 AA; 29918 MW; 093E6EFF2E629CDC8 CRC64;



Query Match	61.48;	Score 836;	DB 11;	Length 258;
Best Local Similarity	61.98;	Pred. No. 8e-45;		
Matches 164; Conservative	36;	Mismatches 57;	Indels 8;	Gaps 3

```
QY      1 MKAAVLTAVLEITGSQARHFWQDEPPQSPWDRVKDLATVYVDYLKDSGRDVSQFEGS 60
        ||||| : : : ||| | | ||||| | ||||| ||||| : ||||| ||||| |
Db      1 MKAAVLAVALVELTGCQAMEWQQDEPSQ - WDRVKDFATVYDAVKDSGRDVSQFESS 58
```

QY 61 ALGKQLNLKLDNWDSDVTSTFSKLRQLGPVTOEFWNLEKETEGLRQEMSKDLEEVRAK 120  
 ||||| |||||: || :||| |||||: || |||||: || ||||| || ||  
 Db 59 TLGKQLNLKLDNWDTLGSTVGRQLQELGPVTOEFWANLEKETDWP RNEMNKDLEENYQK 118

[illegible]

QY	181	DALRTHLAPYSDELRORLAARLEALKENGARLAEYHAKATENHLSLSEKAKPALEDLRO	240
		::   :: :	:: :
Db	175	DALRAKEGLYSDQMRENLAQRLETKNH--PTLIEYHTKASDHILKTGLGEKAKPALDILGQ	232

```
QY 241 GLLPVLESEFKVSFLSALEYTKLN 265
    ||:||||:| :|::| ||||
Db 233 GLMPVLEAWKAKIMSMIDEAKKLN 257
```

RESULT	5
008877	

DT	01-JUL-1997	(TREMBlrel. 04, Created)
DT	01-JUL-1997	(TREMBlrel. 04, Last sequence update)

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OS Rattus norvegicus (Rat).  
 DE NC\_010909.1  
 AF010909.1 A-1.

```

OA      NCBI_TaxID=10110;
RN      [1]
RP      SEQUENCE FROM N.A.

```

Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
 "Repetitive elements in the third intron of murine apolipoprotein A-1"  
 MEDLINE=38011648; PubMed=9415807;

KL	Biochem. Mol. Biol. Int.	43:989-996(1997)
DR	EMBL; U79576; AAB58428.1; -	
DR	HSSP; P02647; 1AV1.	

DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SEQUENCE 258 AA: 29831 MW: 093FR5R2F629CNC8 CRC64.

Query Match	61.28;	Score	834;	DB	11;	Length	258;
Best Local Similarity	61.9%;	Pred.	No.	1	1e-44;		

QY 1 MKA VLT L A V L F L T G S A R H F W Q O D E P P O S P W D R V K D L A T V Y V D V L K D S G R D V S Q F E G S 60

61 ALGQOLNKLIDNWDVSTSTFSKLRQOLGPVIOEFWDNIKEKETEGLRQEMSKDLEEKAK 120

121 VÖPYLDDFÖKKWÖEMELYRÖKVEPLRAELÖEGARÖKTHELÖEKLSPJGEEMDRARAHV 180

181 DALRTHLAPYSDLRQRLARLEALKENGARLAEYHAKATEHLSLSEKAKPAEDLRQ 240

QY 241 GLLPVLESEKVSFLSALLEEYTKKL 265  
 ||:||||: | :|:| | |||  
 Db 233 GLMPVLEAVKAKIMSMIDEAKKKL 257

RESULT	6
Q9TS49	

DT	01-MAY-2000 (TREMBlrel. 13, Created)
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE  
AFOLIOPOLYIN A-1, APOA-1=CHOLESTEROL TRANSPORTER.  
OS  
Erinaceus europaeus (western European hedgehog).  
OC  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;

```
ON  mcepl_1a41d-5503;  
RN  [1]  
RP  SEQUENCE.
```

RT "Plasma lipid transport in the hedgehog: partial characterization of

DR InterPro: IPR000074; Apolipoprotein.  
HSSP: P02647; LAVI.  
DR v. Lipid res. 30;483-495(1995).  
DR v. Lipid res. 30;483-495(1995).

SEQUENCE	241 AA;	2/630 MM;	2EF00F2B69210535	CRC64
30				

Best Local Similarity 59.8%; Pred. No. 4.2e-38;  
Matches 143; Conservative 32; Mismatches 63; Indels 1; Gaps 1;

```
Db 4 KSYWDQIKDMLTYVDATAKDSGKDYLTSLDTSALGQQLNKKLADNWDIVSSALLKAREQM 63
```

```

Db      64 KPIAMEFWGNLEKDEGLRÖTVSKDLELVKEVQPYLDSFÖKVEEELLYRÖKVAPLSA 122
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 124 EWREQARQKAQELQOKAGELGQOHRDRVRTHVDALRTDLAPYGEEARLLQRLQDİKAK 18

```
Db      | 184 SG-DLAEYQIRKISEHILKSFGKEAQPILDRIHGLEPLWEGIKAGAMSMEFLGKKLN$Q 241
```

RESULT 7  
Q9Y355  
ID Q9Y355  
PRELIMINARY: PRT. 67 AA

DT	01-NOV-1999	(TReMBLrel. 12, Created)
DT	01-NOV-1999	(TReMBLrel. 12, Last sequence update)
DT	01-OCT-2000	(TReMBLrel. 15, Last annotation update)

05 Homo sapiens (Human).  
0C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
0C Mammalia; Eutheria; Primates; Carnivora; *Ursidae*;  
0C *Ursus*.

```

RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=00216322. PubMed-10100255.

```

RT	"A novel apolipoprotein A-1 variant, Arg173Pro, associated with
RT	cardiac and cutaneous amyloidosis.";
Riochem	Riochem 257-594-599-1000,
Riochem	Riochem 257-594-599-1000,

DR HSSP; P02647; 1AV1.  
KW Lipoprotein.  
FT NON\_TER 1  
RT NON\_TER 1  
SQ SEQUENCE 67 AA; 7433 MW; 525E1FEB7BDD5AFB CRC64;

Query Match  
Best Local Similarity 24.0%; Score 327; DB 4; Length 67;  
Best Local Similarity 98.5%; Pred. No. 7.4e-14;  
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 165 LSPLGEEMDRARAHVDALRTHLAPYSDELRLARLALKEKNGCARLAETHAKATEHL 224  
DB 1 LSPLGEEMDRARAHVDALRTHLAPYSDELRLARLALKEKNGCARLAETHAKATEHL 60

QY 225 STLSEKA 231  
DB 61 STLSEKA 67

RESULT 8  
Q29248 PRELIMINARY; PRT; 82 AA.  
AC Q29248;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE APOLIPOPROTEIN A-I (FRAGMENT).  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SMALL INTESTINE;  
RX MEDLINE=96327607; PubMed=8672129;  
RA Winteroe A.K., Fredholm M., Davies W.;  
RT "Evaluation and characterization of a porcine small intestine cDNA library."  
RL Mamm. Genome 7:509-517(1996).  
DR EMBL; F14858; CAA23298.1; -.  
KW Lipoprotein.  
FT NON\_TER 82  
SQ SEQUENCE 82 AA; 9168 MW; 24625C65CBFEED8 CRC64;

Query Match  
Best Local Similarity 22.9%; Score 312.5; DB 6; Length 82;  
Matches 63; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 1 MKAAYLTLAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQEGS 60  
DB 1 MKAAYLTLAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQEGS 59

QY 61 ALGKQLNLKLLDNWDSVTSTFESK 83  
DB 60 ALGKHLNLKLLXNDSLGXTFTK 82

RESULT 9  
Q98TG6 PRELIMINARY; PRT; 263 AA.  
AC Q98TG6;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
DE 28KDA-2 APOLIPOPROTEIN.  
OS Anguilla japonica (Japanese eel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillioidei;  
OC Anguillidae; Anguilla.  
OX NCBI\_TaxID=7937;  
RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=21175592; PubMed=11278178;  
RA Kondo H., Kawazoe I., Nakaya M., Kikuchi K., Aida K., Watabe S.;  
RT "The novel sequences of major plasma apolipoproteins in the eel Anguilla japonica."  
RT Biochim. Biophys. Acta 1531:132-142(2001).  
RL Biochim. Biophys. Acta 1531:132-142(2001).  
DR EMBL; AB046203; BAB40960.1; -.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 263 AA; 30087 MW; DB8CB0AF30EBE6B1 CRC64;

Query Match  
Best Local Similarity 21.9%; Score 298.5; DB 13; Length 263;  
Best Local Similarity 28.1%; Pred. No. 2.3e-11;  
Matches 74; Conservative 60; Mismatches 124; Indels 5; Gaps 4;

QY 1 MKAAYLTLAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQEGS 59  
DB 1 MKAAYLTLAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQEGS 59

QY 60 SALGQLNLKLLDNWDSVTSTFESKLEQIGPYTQEFWDNLEKETEGIQEMSKDLEEVKA 119  
DB 60 TEY-KDYKRLSQSLDNIQYIOSASALSPYTDVAVSSQFMELTKMDKIQADVQDKK 118

QY 120 KVQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMDRARAH 179  
DB 119 DIQPKRDELKEVQKHDEYRAKLEPLVKEVTEKHQEMELKTQPVVEDLRARIQVN 178

QY 180 VDALRTHLAPYSDELRLARLALKEKNGCARLAETHAKATEHLSTSEKAPALIEDLR 239  
DB 179 VEETKSKLVPIVEAIRAKLTERLEELRTLAEPYQVEYKDHLSALTDVKKVQG--EDLQ 236

QY 240 QGLLPVLESEKVSFLSALEHYTK 262  
DB 237 SKLKYAEELKTKLVALWESLSQ 259

RESULT 10  
Q93601 PRELIMINARY; PRT; 366 AA.  
AC Q93601;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE APOLIPOPROTEIN AIV.  
GN APOAIV.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98221191; PubMed=9553114;  
RA Steinmetz A., Hermann M., Nimpf J., Aebersold R., Ducret A.,  
RA Weinberg R.B., Schneider W.J.;  
RT "Expression and conservation of apolipoprotein AIV in an avian species."  
RT Species.  
RL J. Biol. Chem. 273:10543-10549(1998).  
DR EMBL; Y16534; CAA76273.1; -.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Lipoprotein.  
SQ SEQUENCE 366 AA; 40852 MW; B47DB49F3E8D91CE CRC64;

Query Match  
Best Local Similarity 18.9%; Score 257.5; DB 13; Length 366;  
Best Local Similarity 26.3%; Pred. No. 1.2e-08;  
Matches 72; Conservative 59; Mismatches 110; Indels 33; Gaps 5;

QY 2 KAAVLTAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQEGSA 61

Db 4 KALL-LVILAVAGTFA-----DVSP-----DQVATVLMRYFTELGSSAKETVDQLOQAE 52  
QY 62 LGKOLNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAKV 121  
Db 53 ITKQNLTLQSLNLSNSYAEELQRLVFPATLQAOVLQDSQRLKQIQOELAELOAKL 112  
QY 122 QPYLDDFQKKQWQEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMDRARAHV 181  
Db 113 APYADEVHQOIGTINRELQAKLSPYADELRSQVDRGTGELRLRALEPPATELREKLQDNAD 172  
QY 182 ALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 241  
Db 173 SIQASLGPYAEIRLQOQIDSSVEGLK-----GQLTPLADELKEQVAQSVEGLRKG 221  
QY 242 LLPV-----LESEKVSFLSALEYTKKL 264  
Db 222 LSPYAEVQDGLNRQLSLTAQMERAAEELRSRL 255

## RESULT 11

001488

ID 001488 PRELIMINARY; PRT; 435 AA.

AC 001488; 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE VARIANT APOLIPOPROTEIN A-IV PRECURSOR (APOA-IV).

GN APOA-4.

OS Mus musculus castaneus (southeastern Asian house mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBITaxID=10091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91286309; PubMed=1648102;

RA Reue K., Leete T.H.;

RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and deletion in a region of tandem repeats.";

RL J. Biol. Chem. 266:12715-12721(1991).

CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN

CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR COMPONENT OF HDL AND CHYLOMICRONS. IT MAY PLAY A ROLE IN THE

CC INTRAVASCULAR METABOLISM OF HDL. POTENTIAL ROLE IN CELLULAR CHOLESTEROL EFFLUX.

CC -1- TISSUE SPECIFICITY: INTESTINE, LIVER, AND PLASMA.

CC -1- MISCELLANEOUS: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A &amp; B, RELATED 11-MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-HELICAL, &amp; MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY

CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH

CC LECITHIN:CHOLESTEROL ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.

CC -1- MISCELLANEOUS: THE APOA-IV POLYMORPHISMS ARE CLASSIFIED BASED ON THE OCCURENCE OF THREE, FOUR, OR FIVE E-Q-A-Q/ E-Q-Q-Q/ E-Q-V-Q REPEAT UNITS.

CC EMBL; M64250; AAA37216.1; -.

DR InterPro; IPR000074; Apolipoprotein.

DR Pfam; PF01442; Apolipoprotein; 1.

KW Lipid transport; HDL; VLDL; Chylomicron; Repeat; Signal; Liver;

KW Plasma; Multigene family; Polymorphism; Cholesterol metabolism.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 435 VARIANT APOLIPOPROTEIN A-IV.

FT DOMAIN 374 393 TANDEM REPEATS.

FT REPEAT 374 377 1.

FT REPEAT 378 381 2.

FT REPEAT 382 385 3.

FT REPEAT 386 389 4.

FT REPEAT 390 393 5.

SQ SEQUENCE 435 AA; 49254 MW; B2AF55EAE8E456B2 CRC64;

Query Match

18.6%; Score 253; DB 11; Length 435;

Best Local Similarity 24.7%; Pred. No. 2.8e-08;  
Matches 66; Conservative 66; Mismatches 103; Indels 32; Gaps 5;  
QY 1 MKAAYTLVAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60  
Db 3 LKAAVLTALVAITGTAK--VTSQVANVWD-----YFTHLSNNAKEAAEQFOKT 52  
QY 61 ALGKOLNLKLLDNWDSVTSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
Db 53 DVTQQLSTLLFQDKLGDASTYADGVRNKLVPFVQVLSGHLAKETERVKEETKELEDLRDR 112  
QY 121 VQPYLDDFQKKQWQEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMDRARAHV 180  
Db 113 MPMHANKVTQTGFENMQKQEHKLPYAVDLQDQINTQTQEMKQLTPYIORMQTIKENV 172  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240  
Db 173 DNHTSMPLATNLKDKFRNMWELK-----GHLPANLAKATIDQN---LEDLRR 221  
QY 241 GLPYLESEKVSFLSALEYTKKLNTQ 267  
Db 222 SLAPLTGVQ-----EKLNHQ 237

## RESULT 12

002762

ID 002762 PRELIMINARY; PRT; 56 AA.

AC 002762; 01-JUL-1997 (Tremblrel. 04, Created)

DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE APOLIPOPROTEIN A1 (FRAGMENT).

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBITaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99098454; PubMed=9883985;

RA Robertson J.A., Bhattacharya S., Ing N.H.;

RT "Tamoxifen up-regulates oestrogen receptor-alpha, c-fos and glyceroldehyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium.";

RL J. Steroid Biochem. Mol. Biol. 67:285-292(1998).

DR EMBL; U94720; AAB57840.1; -.

DR HSP; P02647; IAVL.

KW Lipoprotein.

FT NON\_TER 1 1

FT NON\_TER 56 56

SQ SEQUENCE 56 AA; 6617 MW; 2AB38E08F1E8F1BC CRC64;

## Query Match

18.1%; Score 247; DB 6; Length 56;

Best Local Similarity 82.1%; Pred. No. 5.6e-09;

Matches 46; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 71 LDNWDVSTSTFSKLRQGLPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAKVQPYLD 126

Db 1 LDNWDVSTSTFSKLRQGLPVTQEFWDLNLEKETASLRQEMSKDLEEVKAKVQPYLD 56

## RESULT 13

Q91XF8

ID Q91XF8 PRELIMINARY; PRT; 395 AA.

AC Q91XF8; 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE UNKNOWN (PROTEIN FOR MGC:18592).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBITaxID=10090;





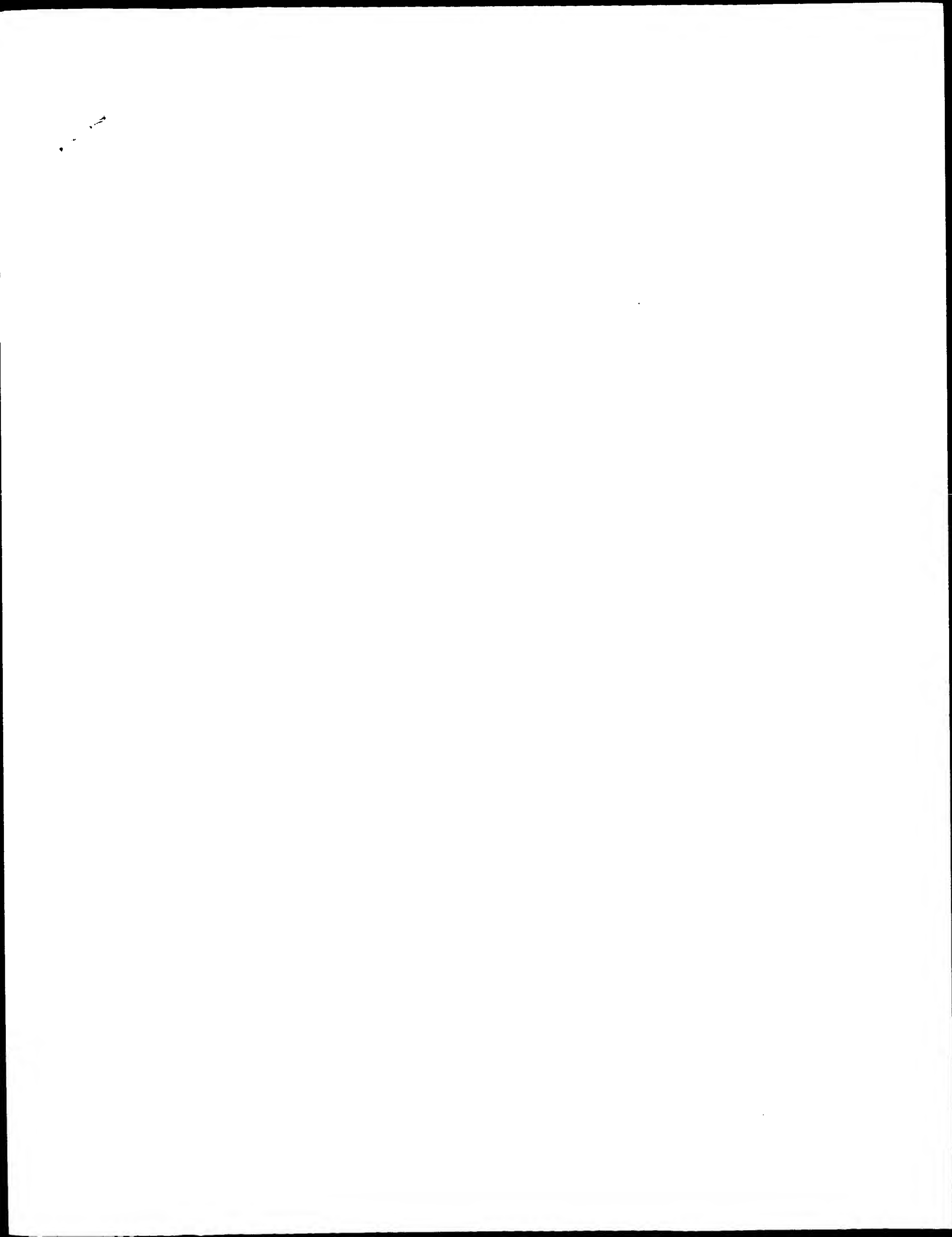
Sun Sep 22 12:10:08 2002

us-09-803-918a-2.rspt

Page 7

[illegible]

Search completed: September 22, 2002, 12:21:36  
Job time: 1116 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:00:25 ; Search time 139.03 Seconds  
(without alignments)  
213.312 Million cell updates/sec

Title: US-09-803-918A-2  
Perfect score: 1362  
Sequence: 1 MRAAVLTAVLFLTGSQARH.....SFKVSPFLSALEEYTKKLTNTQ 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802: \*  
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT: \*  
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT: \*  
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT: \*  
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT: \*  
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT: \*  
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT: \*  
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT: \*  
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT: \*  
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT: \*  
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT: \*  
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT: \*  
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT: \*  
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT: \*  
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT: \*  
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT: \*  
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT: \*  
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT: \*  
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT: \*  
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT: \*  
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: \*  
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT: \*  
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	100.0	267	7 AAP61079	Assumed human apol
2	1362	100.0	267	9 AAP82128	Entire human prepr
3	1362	100.0	267	14 AAR34032	Sequence of apo AI
4	1362	100.0	267	16 AAR72705	Human apo A-I incl
5	1362	100.0	267	20 AAY18675	Human apolipoprote
6	1362	100.0	267	22 AAB47620	Full length Apo-AI
7	1354	99.4	267	18 AAW08602	Human apolipoprote
8	1351.5	99.2	268	9 AAP80668	Recombinant human
9	1328	97.5	299	22 AAU33170	Novel human secret
10	1241	91.1	243	9 AAP81082	Sequence of mature
11	1237.5	90.9	264	15 AAR56863	Apo-lipoprotein AI

12	1237.5	90.9	264	15	AAR56864	Apo-lipoprotein AI
13	1216.5	89.3	244	22	AAU28184	Novel human secret
14	1009	74.1	221	22	AAU29835	Novel human secret
15	912.5	67.0	318	22	AAU30268	Novel human secret
16	831	61.0	359	22	AAU30470	Novel human secret
17	765.5	56.2	166	22	AAU28372	Novel human secret
18	737	54.1	151	22	AAU02278	Human polypeptide
19	617.5	45.3	154	22	AAU12095	Human polypeptide
20	597.5	43.9	120	22	AAU30267	Novel human secret
21	597.5	43.9	120	22	AAU30469	Novel human secret
22	547	40.2	119	22	AAU30468	Novel human secret
23	278.5	20.4	396	22	AAB90664	Human secreted pro
24	269	19.8	382	22	AAB90663	Human secreted pro
25	260.5	19.1	346	14	AAR39493	Human apoAIV mutei
26	259.5	19.1	342	14	AAR39498	Human apoAIV mutei
27	257.5	18.9	337	14	AAR39492	Human apoAIV mutei
28	256.5	18.8	333	14	AAR39497	Human apoAIV mutei
29	255.5	18.8	337	14	AAR39494	Human apoAIV mutei
30	254.5	18.7	342	14	AAR39491	Human apoAIV mutei
31	251.5	18.5	333	14	AAR39490	Human apoAIV mutei
32	251	18.4	391	22	AAB90665	Human secreted pro
33	250.5	18.4	377	14	AAR39489	Human apoAIV mutei
34	249	18.3	372	14	AAR39502	Human apoAIV mutei
35	248.5	18.2	342	14	AAR39487	Human apoAIV mutei
36	248.5	18.2	342	14	AAR39496	Human apoAIV mutei
37	247.5	18.2	333	14	AAR39488	Human apoAIV mutei
38	245.5	18.0	333	14	AAR39481	Human apoAIV mutei
39	245.5	18.0	333	14	AAR39495	Human apoAIV mutei
40	245.5	18.0	377	14	AAR39443	Human apolipoprote
41	245.5	18.0	377	14	AAR45244	Human apoAIV mutei
42	244.5	18.0	377	14	AAR39501	Human apoAIV mutei
43	243.5	17.9	377	14	AAR45242	Human apoAIV mutei
44	241.5	17.7	363	14	AAR39479	Human apoAIV mutei
45	241.5	17.7	373	14	AAR39486	Human apoAIV mutei

ALIGNMENTS

RESULT 1	
AAAP61079	
ID AAP61079 standard; Protein; 267 AA.	
XX	
AC AAP61079;	
XX	
DT 07-OCT-1991 (first entry)	
XX	
DE Assumed human apolipoprotein A-1 derivative gene product.	
XX	
KW Hyperlipaemia; arteriosclerosis.	
XX	
OS Homo sapiens.	
XX	
PN JP61096998-A.	
XX	
PD 15-MAY-1986.	
XX	
PF 16-OCT-1984; 84JP-0216988.	
XX	
PR 16-OCT-1984; 84JP-0216988.	
XX	
PA (MITU ) MITSUBISHI CHEM IND KK.	
XX	
DR WPI; 1986-165025/26.	
DR N-PSDB; AAN60886.	
XX	
PT Human apo-lipoprotein A-1 (deriv.) prepn. - by providing DNA	
PT fragment in cloning site downstream of expression vector promoter	
PT and introducing into host microorganism.	
XX	
PS Disclosure; Fig 2; 9pp; Japanese.	
XX	
CC The human apolipoprotein may be produced by a suitable transformed	





CC contains: (a) a first amino acid residue sequence up to 250 residues  
CC in length that includes residues 120-135 of apo A-I, (b) a second  
CC amino acid residue sequence up to 375 residues in length that  
CC includes residues 217-297 of apo B-100 and DNA encoding it.  
XX  
SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 14; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-99;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEQS 60  
DB 1 mkaavltlavflflltgsqarhfwqdeppspmdrvkdlatvvydvldksgrdyvsqfegs 60  
QY 61 ALGKQLNLKLDNMWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnmwdsvtstfslrqlgptqefwdnlketeqlrgemskdleevkak 120  
QY 121 VQPYLDDEFOKKWQEMELRYQKVEPLRAELQEGARQKLHELOEKLSPGEMRDRARAHV 180  
DB 121 vqpylddfqkkwqeemelyrqqveplraelqegargklheleqeklsplgeemrdrarahv 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapysdelrqlraarleaalkengarlaeyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESFKVSFLSALAEYTKKLTNQ 267  
DB 241 glpvlsefkvsflsaeeytkkltntq 267

RESULT 4  
ID AAR72705 standard; Protein; 267 AA.  
AC AAR72705;  
XX 31-OCT-1995 (first entry)  
DT Human apo A-I including signal and propeptide sequences.  
DE Human apo A-I including signal and propeptide sequences.  
XX Apo A-I; LDL cholesterol; low density lipoprotein; lipid.  
KW Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Peptide /label= presignal  
FT Peptide 19..24  
FT Peptide /label= propeptide  
FT Peptide 120..135  
FT Peptide /label= claimed  
FT Peptide /note= "as part of fusion polypeptide"  
FT Peptide /label= claimed  
FT Peptide /note= "as part of fusion polypeptide"

XX US5408038-A.  
XX 18-APR-1995.  
XX 09-OCT-1991; 91US-0774633.  
XX 09-OCT-1991; 91US-0774633.  
PR 18-JUN-1992; 92US-0901706.  
PR 08-OCT-1992; 92US-0959946.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Curtiss LK, Koduri KR, Smith RS, Wiltzium JL, Young SG;  
XX

DR WPI; 1995-161146/21.  
DR N-PSDB; AAQ89634.

XX New apo:lipoprotein B-100 peptide(s) and fusion peptide(s) - used  
PT in assay systems for detecting LDL and HDL cholesterol levels in  
PT body fluids.  
XX  
PS Claim 10; Fig 2; 41pp; English.

XX AAA89634 and AAR72705 depict the AA sequence of human apo A-I and its  
CC corresp. cDNA, including presignal residues and propeptide  
CC residues, according to Seilhamer et al., DNA 3(4):309 (1984).  
CC A dispersible apo A-I/B-100 fusion polypeptide is claimed which  
CC contains a first AA sequence of apo A-I and that includes at  
CC least AA sequence positions 120-135 (see AAR72606) and which reacts  
CC with pan anti-apo AI antibodies such as: AI-4 ATCC HB8744; AI-7  
CC ATCC HB 8745; AI-9 ATCC HB 8741; AI-10 ATCC HB 9200; AI-11 ATCC  
CC HB 9201; AI-12 ATCC HB 9202; AI-13 ATCC HB 9203; AI-14 ATCC HB  
CC 9204; AI-18 ATCC HB 9507.

XX Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 16; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-99;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEQS 60  
DB 1 mkaavltlavflflltgsqarhfwqdeppspmdrvkdlatvvydvldksgrdyvsqfegs 60  
QY 61 ALGKQLNLKLDNMWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnmwdsvtstfslrqlgptqefwdnlketeqlrgemskdleevkak 120  
QY 121 VQPYLDDEFOKKWQEMELRYQKVEPLRAELQEGARQKLHELOEKLSPGEMRDRARAHV 180  
DB 121 vqpylddfqkkwqeemelyrqqveplraelqegargklheleqeklsplgeemrdrarahv 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapysdelrqlraarleaalkengarlaeyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESFKVSFLSALAEYTKKLTNQ 267  
DB 241 glpvlsefkvsflsaeeytkkltntq 267

RESULT 5  
ID AAY18675 standard; Protein; 267 AA.  
AC AAY18675;  
XX 09-JUL-1999 (first entry)  
DT Human apolipoprotein AI protein sequence.  
DE Human apolipoprotein AI protein sequence.  
XX  
KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.

OS Homo sapiens.  
XX  
XX WO9916409-A2.  
XX 08-APR-1999.  
XX 28-SEP-1998; 98WO-US20329.  
XX 29-SEP-1997; 97US-0940136.  
PR  
XX (BUTT/) BUTTNER K.  
PA



CC (AFTI). These fragments are selected from an 18 kD N-terminal fragment  
CC (amino acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144)  
CC and a 13 kD C-terminal fragment (amino acids 156-267). The AFTI  
CC polypeptides and polynucleotides are useful for regulating T-cell  
CC mediated activation of monocytes and for treating, diagnosing,  
CC ameliorating diseases associated with IL-1 and/or TNF activity.  
CC The diseases are acute pancreatitis, Alzheimer's disease, asthma,  
CC cancer, fever, inflammatory bowel disease, ischemia, multiple  
CC sclerosis, osteoporosis, Parkinson's disease or psoriasis. Numerous  
CC examples of other diseases are given in the specification.  
CC The AFTI nucleic acids are useful as hybridization probes in diagnostic  
CC assays to test for the presence of an AFTI or Apo-A-I DNA in mammalian  
CC tissue or bodily fluid samples.

XX Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 22; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-99;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTLAVLFLTGSAHFHWQDEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEQS 60  
DB 1 mkaavltlavflitgsqarhfwqdeppspwdrvkdlatvyvdvlkdsgrdyvsqfegs 60  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQIGPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnwdsvststfsklreqlpgvtqefwldnleketeglrqemskdleevkak 120  
QY 121 VQPYLDLDFQKQWQEMELYRQKVEPLRAELQEGAROKLHELOEKLSPLGEMRDRARAHV 180  
DB 121 vqpyldldfqqkwqeemelyrqkveplraelqegarklhelqeklsplgeemrdrarahv 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapydsdelrqlraarleaalkengarlaeyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESEFKVSFLSALEEYTKKLTNQ 267  
DB 241 gllpvlesfkvsflsaleeytkkltntq 267

RESULT 7  
AAW08602 standard; Protein; 267 AA.  
XX AAW08602;  
AC AAW08602;  
XX 04-SEP-1997 (first entry)

DE Human apolipoprotein A-1 variant "Paris" protein sequence.

KW Human; apolipoprotein A-1; variant; mutation; serum lipid; cholesterol;  
KW high density lipoprotein; triglyceride; symptom; cardiovascular disease;  
KW atherosclerosis; restenosis; myocardial infarct; angina; death; stroke;  
KW cardiac decompensation; metabolic deficit.

XX Homo sapiens.

OS Homo sapiens.

FT Key Location/Qualifiers  
FT Peptide 1..24  
FT Region /note= "prepro peptide"  
FT 172..178  
FT /note= "this sequence which contains the mutated amino  
FT Misc-difference 175 acid residue is claimed - claim 2"

FT /note= "changed from Arg residue in wild type protein,  
FT due to a C to T transition mutation"

PN WO9637608-A1.

XX 28-NOV-1996.

PF 20-MAY-1996; 96WO-FR00747.  
XX 22-MAY-1995; 95FR-0006061.  
PR (INSP ) INST PASTEUR LILLE.  
PA (RHON ) RHONE POULENC ROBER SA.  
PA (UYPA-) UNIV CURIE PARIS VI P & M.  
XX Benoit P, Bruckert E, Denefle P, Duverger N, Fruchart J;  
PI Luc G, Turping, Assmann G, Funke H;  
XX WPI; 1997-021218/02.  
DR N-PSDB; AAT43691.

PT New variant of human apoA-1 with Cys at position 151 - has  
PT anti-atherogenic activity for treatment and prevention of  
PT cardiovascular disease  
PS Claim 2; Page -: 58pp; French.

CC This is the amino acid sequence of a human apolipoprotein A-1 variant  
CC designated the "Paris" variant which has a Cys replacing the Arg residue  
CC at position 151. The substitution is generated by a mutation of  
CC the C nucleotide at position 523 in the wild type gene to a T residue,  
CC changing the encoded residue from an Arg to a Cys. The gene was isolated  
CC from a patient with an unusual pattern of serum lipids i.e. low levels of  
CC apoA-1 and high density lipoprotein (HDL)-cholesterol and high  
CC triglycerides, but showing symptoms of atherosclerosis. The new variant  
CC protein has anti-atherogenic activity so is useful for treatment and  
CC prevention of cardiovascular diseases such as atherosclerosis,  
CC restenosis, myocardial infarct, angina, sudden death, stroke and cardiac  
CC decompensation, or more generally any condition involving genetic or  
CC metabolic deficit of apoA-1.  
CC Note: this sequence is not given in the specification but is generated  
CC from the wild type apoA-1 gene disclosed in the specification and has  
CC the appropriate amino acid changed.

XX Sequence 267 AA;

Query Match 99.48%; Score 1354; DB 18; Length 267;  
Best Local Similarity 99.6%; Pred. No. 7.3e-99;  
Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAAYLTLAVLFLTGSAHFHWQDEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEQS 60  
DB 1 mkaavltlavflitgsqarhfwqdeppspwdrvkdlatvyvdvlkdsgrdyvsqfegs 60  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQIGPVTQEFWDLNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 algkqlnlkldnwdsvststfsklreqlpgvtqefwldnleketeglrqemskdleevkak 120  
QY 121 VQPYLDLDFQKQWQEMELYRQKVEPLRAELQEGAROKLHELOEKLSPLGEMRDRARAHV 180  
DB 121 vqpyldldfqqkwqeemelyrqkveplraelqegarklhelqeklsplgeemrdrarahv 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 dalrthlapydsdelrqlraarleaalkengarlaeyhakatehlstlsekakpaledlrq 240  
QY 241 GLLPVLESEFKVSFLSALEEYTKKLTNQ 267  
DB 241 gllpvlesfkvsflsaleeytkkltntq 267

RESULT 8  
AAP80668 standard; protein; 268 AA.  
XX AAP80668;  
AC AAP80668;  
XX 24-OCT-1990 (first entry)

DE Recombinant human preproapoprotein A1.  
XX human preproapoprotein A1; high density lipoprotein deficiency; ss.  
KW synthetic.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Peptide /label=precursor  
FT Peptide 19..25  
FT Protein /label=proprotein  
FT Protein 26..268  
FT Protein /label=mature apoprotein A1  
XX  
XX EP293357-A.  
XX  
XX 30-NOV-1988.  
XX  
XX 24-MAY-1988; 88EP-0870095.  
XX  
XX 28-MAY-1987; 87GB-0012540.  
XX  
XX (UNIO ) UCB SA.  
XX  
XX Bollen A, Gobert J, Wulfert E;  
XX  
XX WPI; 1988-339891/48.  
XX  
XX N-PSDB; AAN81258.  
XX  
XX  
PT New DNA encoding human preproapoprotein A1 -  
PT modified to eliminate hairpin structures  
XX  
PS Claim 1; Page 12; 25pp; French.  
XX  
XX Met at posn 19 is inserted as an extra amino acid c.f. wild-type  
XX  
XX protein.  
XX The DNA fragment used to replace the wild-type sequence encoding  
XX amino acids -6 to +14 (= Arg 20 to Leu 39) modifies codons  
XX corresponding to wild-type amino acids -6, -1, 1, 3, 4, 5, 6, 7,  
XX 10, 11 and 14. The changed codons still encode the same amino acids  
XX as in the wild-type protein but reduce formation of secondary  
XX structures in mRNA.  
XX See also AAN82064.  
XX  
XX Sequence 268 AA;  
SQ  
Query Match 99.2%; Score 1351.5; DB 9; Length 268;  
Best Local Similarity 99.6%; Pred. No. 1.2e-98;  
Matches 267; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 MKAAYLTIAVLFTGSQA-RHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEG 59  
DB 1 mkaavltlavltfsgqamrhfwqdeppqspwdrvkdlatvyvdlkdsgrdyvsqfeg 60  
QY 60 SALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQGFWDNLKETEGLRQEMSKDLEEVKA 119  
DB 61 salgqnlklnldnwdsvtstfsklreqlqpvrtqefwlnleketeglrqemskdleevka 120  
QY 120 KVPYLDLDFQKKWQEMELYRQKVEPLRAELQEGARQKLHLEQEKLSPLGEMDRARAH 179  
DB 121 kvpyldldfqqkwqemelyrqkveplraelqegarqklhelqeklsplgeemdrarah 180  
QY 180 VDALRTHLAPYSDELRLQRLAARLEALKENGARLALEYHAKATEHLSLSEKAKPALDLR 239  
DB 181 vdalrthlapysdelrlqrlaarlealkengarlaeayahakatehlslsekakpaledlr 240  
QY 240 QGLLPVLESFVKVSLSALEEYTKKLNQ 267  
DB 241 qglipvlesfkvsfisaaleeytkklnq 268  
RESULT 9

AAU33170  
ID AAU33170 standard; Protein; 299 AA.  
XX  
XX AC AAU33170;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
XX DE Novel human secreted protein #3661.  
XX  
XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX WO200179449-A2.  
XX  
XX 25-OCT-2001.  
XX  
XX 16-APR-2001; 2001WO-US08656.  
XX  
XX 18-APR-2000; 2000US-0552929.  
XX  
XX 26-JAN-2001; 2001US-0770160.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-611725/70.  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
XX  
PS Claim 20; Page 718; 765pp; English.  
XX  
XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising  
XX the nucleic acids encoding the polypeptides and cells genetically  
XX engineered to express them are also useful for producing the proteins.  
XX The proteins are useful in genetic vaccination, testing and  
XX therapy, and can be used as nutritional supplements. They may be used to  
XX increase stem cell proliferation; to regulate haematopoiesis; and in  
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
XX immune suppression and/or stimulation; as anti-inflammatory agents; and  
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
XX sequences of novel human secreted proteins of the invention.  
XX  
XX Sequence 299 AA;  
SQ  
Query Match 97.5%; Score 1328; DB 22; Length 299;  
Best Local Similarity 97.0%; Pred. No. 9.4e-97;  
Matches 259; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKAAYLTIAVLFTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60  
DB 33 mkaavltlavltfsgqarhfwqdeppqspwdrvkdlatvyvdlkdsgrdyvsqfegs 92  
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQGFWDNLKETEGLRQEMSKDLEEVKA 120  
DB 93 algkqlnlklnldnwdsvtstfsklreqlqpvrtqefwlnleketeglrqemskdleevka 152  
QY 121 VQPYLDLDFQKKWQEMELYRQKVEPLRAELQEGARQKLHLEQEKLSPLGEMDRARAHV 180  
DB 153 vqpyldldfqqkwqemelyrqkveplraelqegarqklhelqeklsplgeemdrarahv 212  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLALEYHAKATEHLSLSEKAKPALDLRQ 240



Db 213 dalrthlapysgelrqlgarlgalrengarmgqyhaqatehltstlsekakpaledlrg 272  
QY 241 GLLPVLESFVKVSLSALEEYTKLNTQ 267  
273 glpvlsefksfslsaleeytklntq 299

## RESULT 10

AAP81082  
ID AAP81082 standard; protein; 243 AA.  
XX  
AC AAP81082;  
XX

DT 14-JAN-1991 (first entry)  
XX

DE Sequence of mature human apolipoprotein AI (apoAI).  
XX

KW Atherosclerosis; therapy; cardiovascular disease.  
XX

OS Homo sapiens.  
XX

PN WO8803166-A.  
XX

PD 05-MAY-1988.  
XX

PF 21-OCT-1987; 87WO-EP00621.  
XX

PR 23-OCT-1986; 86GB-0025435.  
XX

PA (FARM ) FARMITALIA C ERBA SPA.  
XX

PI Lorenzetti R, Monaco L, Soria M, Palomba R, Isacchi A, Sarmientos P;  
XX

DR WPI; 1988-133240/19.  
XX

DR N-PSDB; AAN80243.  
XX

PT Recombinant human apo:lipoprotein AI -  
XX

PT used to lower plasma cholesterol and/or tri glyceride levels and  
XX

PT to combat atherosclerosis and cardiovascular diseases  
XX

PS Disclosure; Fig 1; 51pp; English.  
XX

CC The protein comprising apo AI genetic variants may be used to lower  
XX

CC plasma cholesterol and/or triglyceride levels. They may also be used to  
XX

CC combat atherosclerosis and cardiovascular diseases such as coronary  
XX

CC heart disease. Prefd. proteins are Met-apo AI, Met-apo AI-T6, Met-apo  
XX

CC AI-MI and Met-apo AI-T6/MI.  
XX

CC Sequence 243 AA;  
XX

Query Match 91.1%; Score 1241; DB 9; Length 243;  
XX

Best Local Similarity 100.0%; Pred. No. 5.2e-90;  
XX

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX

QY 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLIDNWDVSTFSKL 84  
XX

Db 1 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlidnwdsvstfskl 60  
XX

QY 85 REQLGPTVQEFWDLNLEKETEGLRQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELRYQKVE 144  
XX

Db 61 reglqpvteqfwdlnleketeglrqemskdleevkakvpylddfqkkwqeemelyrqkve 120  
XX

QY 145 PLRAELQEGARQKLHELQEKLSPIGEMDRARAHVDAIRTHLAPYSDELRORLARLEA 204  
XX

Db 121 plraelqegarklhelqeklspligemdrarahvdairthlapysdelrqlarlea 180  
XX

QY 205 LKENGARLAIFYHAKATEHLSTLSEKAKPALEDLROGLPVLSEFVKVSLSALEEYTKKL 264  
XX

Db 181 lkengarlaifyhakatehltstlsekakpaledlrglplvlesfksfslsaleeytkkl 240  
XX

Db 241 ntq 243

## RESULT 11

AAR56863  
ID AAR56863 standard; Protein; 264 AA.  
XX

AC AAR56863;  
XX

DT 26-JAN-1995 (first entry)  
XX

DE Apo-lipoprotein AI-M.  
XX

KW Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid pKP683;  
XX

KW vector; atherosclerosis; cardiovascular disease.  
XX

OS Homo sapiens.  
XX

PN WO9413819-A.  
XX

PD 23-JUN-1994.  
XX

PF 09-DEC-1993; 93WO-SE01061.  
XX

PR 11-DEC-1992; 92SE-0003753.  
XX

PA (KABI ) KABI PHARMACIA AB.  
XX

PI Abrahmsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
XX

PI Sejltz T;  
XX

DR WPI; 1994-217892/26.  
XX

DR N-PSDB; AAO68357.  
XX

PT Expression vector for extracellular prodn of apo-lipoprotein AI-M  
XX

PT in E. coli - includes inducible promoter and signal peptide  
XX

PT sequences providing improved yield, useful in treating  
XX

PT atherosclerosis and cardiovascular disease.  
XX

PS Disclosure; Fig. 3; 33pp; English.  
XX

CC Plasmid pKP683 encodes human apo-lipoprotein AI-M in E. coli. The  
XX

CC NotI-HindIII segment of pKP683 and the deduced amino acid sequence  
XX

CC of the translated Apo AI-M protein were determined. The plasmid  
XX

CC provides high yields of extracellular Apo AI-M.  
XX

CC Sequence 264 AA;  
XX

Query Match 90.9%; Score 1237.5; DB 15; Length 264;  
XX

Best Local Similarity 91.9%; Pred. No. 1.1e-89;  
XX

Matches 248; Conservative 4; Mismatches 9; Indels 9; Gaps 2;  
XX

QY 1 MKAAYLTFLAVL---FLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQF 57  
XX

Db 1 mktaiaiavalagfatvana-----deppqspwdrvkdlatvyvdvlkdsgrdyvsqf 54  
XX

QY 58 EGSAIGKQLNLKLDNWDVSTFSKLREQLGPTVQEFWDLNLEKETEGLRQEMSKDLEEV 117  
XX

Db 55 egasalqkqlnlkldnwdsvstfsklreglqpvteqfwdlnleketeglrqemskdleev 114  
XX

QY 118 KAKVQPYLDLDFQKKWQEMELRYQKVEPLRAELQEGARQKLHELQEKLSPIGEMDRAR 177  
XX

Db 115 kakvpylddfqkkwqeemelyrqkveplraelqegarklhelqeklspligemdrar 174  
XX

QY 178 AHVDALRTHLAPYSDELRORLAARLEALKENGARLAIFYHAKATEHLSTLSEKAKPALED 237  
XX

Db 175 ahvdalrthlapysdelrqlaarlealkengarlaifyhakatehltstlsekakpaled 234  
XX

QY 238 LROGLPVLSEFVKVSLSALEEYTKLNTQ 267  
XX

Db 235 lrglplvlesfksfslsaleeytklntq 264  
XX



## RESULT 12

AAAC0004  
ID AAR56864 standard; Protein; 264 AA.

AC AAR56864;

DT 26-JAN-1995 (first entry)

Apo-lipoprotein AI-M. DE

AA Apo-lipoprotein AI-M; Apo AI-M; Escherichia coli; plasmid pKF764;  
KM KM vector; atherosclerosis; cardiovascular disease.  
KM vector; atherosclerosis; cardiovascular disease.

OS Homo sapiens.

PN W09413819-A.

PD 23-JUN-1994

PF 09-DEC-1993; 93WO-SE01061.

PR 11-DEC-1992; 92SE-0003753.

PA (KABI ) KABI PHARMACIA AB.

PI Abrahamsen L, Holmgren E, Kalderen C, Lake M, Mikaelsson A;  
PI Sejlitz T;

DR WPI; 1994-217892/26.  
DR N-PSDB; AAQ68358.

PT Expression vector

aa Expression vector for extracellular prodn of apo-lipoprotein AI-M  
PT in E. coli - includes inducible promoter and signal peptide  
PT sequences providing improved yield, useful in treating  
PT atherosclerosis and cardiovascular disease.

PS Disclosure; Fig. 4; 33pp; English.

Plasmid pkr764 encodes human apo-lipoprotein AI-M in *E. coli*. The NotI-HindIII segment of pkr764 and the deduced amino acid sequence of the translated Apo AI-M protein were determined. The plasmid provides high yields of extracellular Apo AI-M.

Sequence	264 AA;
SQ	

Query Match	90.98;	Score 1237.5;	DB 15;	Length 264;
-------------	--------	---------------	--------	-------------

Best Local Similarity	91.98;	Pred. No. 1.1e-89;
Matches 248; Conservative	4;	Mismatches 9; Indels 9; Gaps 2;

OY	1	MKAAVLTIAVL--ELTGSQARHFWMQODEPPQSPMDRKYDLATVYVDVLKDSGRDYSQF	57
Dd	1	mktataiaavalaqafatvana-----deppqspwdrvkdlatvyvdvlkdsgrdyvsqf	54
OY	58	EGSALGKQLNLKLIDNMDSVTSTFSKLRQOLGPVTOEFWDNLEKETEGLRQEMSKDIEEV	117
Dd	55	egsalgkqlnlklidnmdsvstfsklreqlgpvtgeifwnleketeglrgemskdleev	114
OY	118	KAKVPYLDDFQKKWQEEMELYROKVEPLRAELOEGAROKLHELQEKLSPLGEEMDRBR	177
Dd	115	kakvpylddfkkkwqeemelyrqkveplraelqegarqkhelqeklsplgeemdrar	174
OY	178	AHVDAIRTHLAPYSDELRORLARLLALKENGARGARLAETHAKATEHLSTLSEKAKPALED	237
Dd	175	ahvdairthlapysdelrqlaarleaalkengargarlaeyhakatehlstlsekapaled	234
OY	238	LROGLLPVLESFKVSFLSALEEYTKKLNTO	267
Dd	235	lrqglipvlesfkvsflsaaleeytkkln tq	264

## RESULT 13

ID    AAU28184   standard; Protein; 244 AA.

AC AAU28184;

DT 18-DEC-2001 (first entry)

DE Novel human secretory protein, Seq ID No 353.

AA Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
 KW transgenic animal; Alzheimer's disease; parkinson's disease; burn;  
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
 KW fertility; analgesic; pain; antigen.

OS Homo sapiens.

PN WO200166689-A2.

PD 13-SEP-2001.

PF 05-MAR-2001; 2001WO-US04942.

PR 07-MAR-2000; 2000US-0519705.

PR 17-JUN-2000; 2000US-0596193  
DT 14-JUN-2000; 2000US-0616847

PR 19-SEP-2000; 2000US-0665363  
20-OCT-2000; 2000US-0683267

XX  
XX  
/HYGE-) HYSGO INC

XX  
XX  
FT

Farrar VT  
Till C.  
Asundi V.

PI Zhao QA, yang x, Dillmanac  
xy

DR WPL; 2001-589934/00.  
DP N-PSTDB: AAS45084

XX	Novel polynucleotide
XX	
PT	

PT prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders -

Example 5: SEO ID No 353: 107bp; English  
XX  
PS

The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis,

CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
CC reactions and conditions, such as asthma or other respiratory problems.  
CC In addition, (1) affects biorhythms or circadian cycles of rhythms,  
CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
CC analgesic effects or other pain reducing effects, immunoglobulin like  
CC activity and can act as an antigen in a vaccine composition to raise an  
CC immune response. AAU28020-AAU28395 represent novel human secreted protein  
CC amino acid sequences of the invention.  
XX  
SQ Sequence 244 AA;

Query Match 89.3%; Score 1216.5; DB 22; Length 244;  
Best Local Similarity 91.0%; Pred. No. 4.4e-88;  
Matches 243; Conservative 1; Mismatches 0; Indels 23; Gaps 1;

QY 1 MKAAYLTAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFECS 60  
Db 1 mkaavltavfltgsgarhfwqdeppqspwdrvkdlatvyvdlkdsq----- 51  
QY 61 ALGKQLNLKLDNWDSTSTFSKLRQGLPVTOEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
Db 52 -----dsvtstfsklrqlgptvqgfdwnleketeglrqemskdleevkak 97  
QY 121 VQPYLDDFOKKWQEMELRYRQKVEPLRAELQEGARQKLEHLEQKLSPLGEEMDRARAHV 180  
Db 98 vqpylddfqkwwqeemelyrqkveplraelqegarklhelqeklsplgeemdrarahv 157  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAPALEDLRO 240  
Db 158 dalrthlapydsdelrqlraarlealkengarlarlaeyhakatehlstlsekapaledlrq 217  
QY 241 GLPYLESFQVSLFSALEEYTKKINTQ 267  
Db 218 glpylesfqsflsaeeytkkintq 244

RESULT 14

AAU29835  
ID AAU29835 standard; Protein; 221 AA.

XX AC AAU29835;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #326.

XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 199; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 221 AA;

Query Match 74.1%; Score 1009; DB 22; Length 221;  
Best Local Similarity 87.9%; Pred. No. 8.6e-72;  
Matches 203; Conservative 2; Mismatches 2; Indels 24; Gaps 2;

QY 1 MKAAYLTAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFECS 59  
Db 14 mkaavltavfltgsgarhfwqdeppqspwdrvkdlatvyvdlkdsq----- 65  
QY 60 SALGKQLNLKLDNWDSTSTFSKLRQGLPVTOEFWDNLEKETEGLRQEMSKDLEEVKA 119  
Db 66 -----dsvtstfsklrqlgptvqgfdwnleketeglrqemskdleevka 110  
QY 120 KQPYLDDFOKKWQEMELRYRQKVEPLRAELQEGARQKLEHLEQKLSPLGEEMDRARAH 179  
Db 111 kvpylddfqkwwqeemelyrqkveplraelqegarklhelqeklsplgeemdrarah 170  
QY 180 VDLRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSTLSEK 230  
Db 171 vdlrthlapydsdelrqlraarlealkengarlarlaeyhakatehlstlsek 221

RESULT 15

AAU30268  
ID AAU30268 standard; Protein; 318 AA.

XX AC AAU30268;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #759.

XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
XX  
PS Claim 20; Page 270; 765pp; English.

XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 318 AA;

Query Match 67.0%; Score 912.5; DB 22; Length 318;  
Best Local Similarity 73.9%; Pred. No. 5.4e-64;  
Matches 207; Conservative 17; Mismatches 43; Indels 13; Gaps 13;

QY 1 MKAVALTLAVLELTGSOARHFWQ-QDEPPQSPWDRYKDLAT-VYVDVLKDSGRD-YVSQF 57  
Db 14 mkaavltlavlflltsgarhfwgqxspraawdrvkdlatrvtvlkeqrletyvsqf 73  
QY 58 EGSALGK-QLNLKLLDNW-DSVTSTPS-KLREQLGPVTQEFWDNLEKETEGLRQEMSKDL 114  
Db 74 eglrlgensxtlkildnwgrxpsltfqpscakqlgpltfqefwynleketegfrqemskdl 133  
QY 115 EEVKAKVQPY-LDDEQOKK-WQEEEMELYRQKVEPL-RAELQEGAR-QKLEHL-QEKLSPUG 169  
Db 134 eevkakvqpyliddfgerwqemelyrqkveplarknfdqearpeslhelarrisplg 193  
QY 170 EEM-RDRARAHVDALRTHLAPYSDELRORLA-ARLEALKENGARLAELYHAKATEHLSTL 227  
Db 194 eavsrprarpmwdalrthlapysdemmpalgraplgairengarmgqyhaqatehlstl 253  
QY 228 SEKAKPALEDLRQGLLPVLESEKVSFLSALEEYTKKLNTQ 267  
Db 254 sekakpaledlrqglpvlsefksflsaleeytkklntq 293

Search completed: September 22, 2002, 12:05:15  
Job time: 290 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:06:29 ; Search time 66.91 Seconds  
(without alignments)  
127.813 Million cell updates/sec

Title: US-09-803-918A-2\_COPY\_25\_113

Perfect score: 468  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....EFWDNLEKETEGLRQEMSKD 89

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	468	100.0	267	1	LPHUA1 apolipoprotein A-I
2	443	94.7	267	1	A26529 apolipoprotein A-I
3	443	94.7	267	2	JS0079 apolipoprotein A-I
4	407	87.0	266	1	LPDGA1 apolipoprotein A-I
5	405	86.5	264	2	S31394 apolipoprotein A-I
6	405	86.5	265	2	A46018 apolipoprotein A-I
7	393	84.0	265	2	JT0672 apolipoprotein A-I
8	383	81.8	265	1	LPRB1B apolipoprotein A-I
9	383	81.8	265	2	A56858 apolipoprotein A-I
10	383	81.8	266	1	LPRB1Z apolipoprotein A-I
11	378.5	80.9	241	2	A24998 apolipoprotein A-I
12	359	76.7	259	2	A24700 apolipoprotein A-I
13	355	75.9	231	2	JQ0704 apolipoprotein A-I
14	343	73.3	262	2	JC1237 apolipoprotein A-I
15	343	73.3	264	2	S22420 apolipoprotein A-I
16	223	47.6	264	1	LPCHA1 apolipoprotein A-I
17	223	47.6	264	2	JC5456 apolipoprotein A-I
18	220	47.0	246	2	A61448 apolipoprotein A-I
19	106	22.6	34	2	S67972 apolipoprotein A-I
20	106	22.6	36	2	A56866 apolipoprotein A-I
21	101	21.6	20	2	A05313 apolipoprotein A-I
22	92	19.7	396	1	LPHUA4 apolipoprotein A-I
23	90	19.2	391	1	LPRTA4 apolipoprotein A-I
24	90	19.2	429	2	S29565 apolipoprotein A-I
25	86	18.4	401	2	A47141 apolipoprotein A-I
26	83	17.7	312	1	LPRTA apolipoprotein A-I
27	81	17.3	365	2	D71559 probable ABC ATPase
28	81	17.3	391	2	B40892 apolipoprotein A-I
29	81	17.3	394	2	A25281 apolipoprotein A-I

30	81	17.3	395	2	A40892 apolipoprotein A-I
31	81	17.3	399	2	C40892 apolipoprotein A-I
32	78.5	16.8	2697	2	T25444 hypothetical prote
33	78	16.7	311	2	JU0036 apolipoprotein E p
34	76.5	16.3	207	2	S56209 probable membrane
35	73	15.6	491	2	T27661 hypothetical prote
36	72.5	15.5	164	2	B37842 hypothetical prote
37	72.5	15.5	192	2	AF1987 hypothetical prote
38	71.5	15.3	625	2	G90570 UDP-glucose 4-epim
39	70	15.0	328	2	A84167 tryptophanase (EC
40	70	15.0	453	2	A49022 tryptophanase (EC
41	70	15.0	453	2	B49022 DNA-directed RNA p
42	69	14.7	926	2	H84415 probable tail comp
43	69	14.7	1560	2	T09202 hypothetical prote
44	68.5	14.6	865	2	AC1966 dynein heavy chain
45	68.5	14.6	4725	1	A44357

## ALIGNMENTS

RESULT 1  
LPHUA1  
apolipoprotein A-I precursor [validated] - human  
N:Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C:Species: Homo sapiens (man)  
C>Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C/Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010;  
6197  
R;Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A/Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A/Reference number: A90947; MUID:8502665  
A/Accession: A90947  
A/Molecule type: DNA  
A/Residues: 1-267 <SEI>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
A/Accession: B90947  
A/Molecule type: mRNA  
A/Residues: 1-267 <SE2>  
A/Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R;Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988  
A/Title: Sequence and expression of Tangier apoA-I gene.  
A/Reference number: S02373; MUID:88196137  
A/Accession: S02373  
A/Molecule type: DNA  
A/Residues: 1-267 <MAK>  
A/Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
R;Shoulders, C.C.; Kornblith, A.R.; Munro, B.S.; Baralle, F.E.  
Nucleic Acids Res. 11, 2827-2837, 1983  
A/Title: Gene structure of human apolipoprotein A1.  
A/Reference number: A93465; MUID:83220822  
A/Accession: A93465  
A/Molecule type: DNA  
A/Residues: 1-267 <SHO>  
A/Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518;  
R;Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
A/Title: Isolation and characterization of the human apolipoprotein A-I gene.  
A/Reference number: A21147; MUID:84016011  
A/Accession: A21147  
A/Molecule type: DNA  
A/Residues: 1-267 <KAR>  
A/Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.  
Nucleic Acids Res. 12, 3917-3932, 1984  
A/Title: Human apolipoproteins A1, AII, CII and CIII. cDNA sequences and mRNA abundan  
A/Reference number: A93519; MUID:84221405  
A/Accession: A93519  
A/Molecule type: mRNA  
A/Residues: 1-267 <SHA>  
A/Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753



A:Accession: B93519  
 A:Molecule type: DNA  
 A:Residues: 1-24 <SH2>  
 R:Cheung, P.; Chan, L.  
 Nucleic Acids Res. 11, 3703-3715, 1983  
 A:Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
 A:Reference number: A93472; MUID:83220772  
 A:Accession: A93472  
 A:Molecule type: mRNA  
 A:Residues: 1-267 <CHE>  
 A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:J00102  
 R:Law, S.W.; Brewer Jr., H.B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
 A:Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
 A:Reference number: A94010; MUID:84119464  
 A:Accession: A94010  
 A:Molecule type: mRNA  
 A:Residues: 1-267 <LAW>  
 A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:J00102  
 R:Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
 A:Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
 A:Reference number: A21118; MUID:83195100  
 A:Accession: A21118  
 A:Molecule type: mRNA  
 A:Residues: 1-24 <ZAN>  
 R:Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
 Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
 A:Title: Human plasma proapoA-I: isolation and amino-terminal sequence.  
 A:Reference number: A90112; MUID:83256553  
 A:Accession: A90112  
 A:Molecule type: protein  
 A:Residues: 19-27 <BRE>  
 R:Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
 Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
 A:Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high density lipoproteins  
 A:Reference number: A90209; MUID:78123731  
 A:Accession: A90209  
 A:Molecule type: protein  
 A:Residues: 25-57, 'Q', '59-169, 'QQ', '172-267 <BR2>  
 R:Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
 J. Clin. Invest. 82, 803-807, 1988  
 A:Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A-I)  
 A:Reference number: A30516; MUID:88331387  
 A:Accession: A30516  
 A:Molecule type: protein  
 A:Residues: 25-56 <YUI>  
 R:Nichols, W.C.; Dwulet, F.E.; Liepnieks, J.; Benson, M.D.  
 Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
 A:Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid.  
 A:Reference number: A31582; MUID:89050104  
 A:Accession: A31582  
 A:Molecule type: protein  
 A:Residues: 25-49, 'R', '51-85, 'D', '87-107 <NIC>  
 A:Note: Variant sequence from patient with familial amyloidotic polyneuropathy type III  
 R:Manjunath, P.; Marcel, Y.L.; Uma, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A.  
 J. Biol. Chem. 264, 16853-16857, 1989  
 A:Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
 A:Reference number: A34409; MUID:89380318  
 A:Accession: A34409  
 A:Molecule type: protein  
 A:Residues: 25-48 <MAN>  
 R:Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
 A:Title: Structural requirements of human preproapolipoprotein AI for translocation and secretion  
 A:Reference number: S02737; MUID:89149957  
 A:Accession: S02737  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-6, 'AV', '9, 'LV', '12-29 <STO>  
 A:Note: Part of this sequence, including the amino end of the mature protein, was confirmed by Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A:Title: Transient expression of wild type and mutant human apolipoprotein AI in COS cells  
 A:Reference number: S16197; MUID:92029676  
 A:Contents: annotation; extension of studies in reference S02737  
 R:Stoffel, W.; Kruger, E.; Deutzmann, R.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
 A:Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processed in the rough microsomal fraction  
 A:Reference number: A19913; MUID:83236195  
 A:Accession: B19913  
 A:Molecule type: protein  
 A:Residues: 1-6, 'X', '8-13, 'xxx', '17-18, 'xx', '21, 'X', '23-25, 'X', '27-29 <ST2>  
 R:Ehnholm, C.; Bozas, S.E.; Tenkanen, H.; Kirschaum, L.; Metsu, J.; Murphy, B.; Walke  
 Biochim. Biophys. Acta 1086, 255-260, 1991  
 A:Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
 A:Reference number: A56815; MUID:92075698  
 A:Accession: A56815  
 A:Molecule type: protein  
 A:Residues: 25-31, 'p', '33 <EHN>  
 A:Experimental source: serum  
 A:Note: sequence extracted from NCBI backbone (NCBIP:69759)  
 A:Note: 32-Tip was also found  
 R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
 Biochemistry 33, 1988-1993, 1994  
 A:Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins  
 A:Reference number: A54223; MUID:94162201  
 A:Accession: A54223  
 A:Molecule type: protein  
 A:Residues: 25-39 <KUN>  
 R:Mogulilevsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.;  
 DNA 8, 429-436, 1989  
 A:Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: pAPOL  
 A:Reference number: I39476; MUID:89377481  
 A:Accession: I39476  
 A:Molecule type: mRNA  
 A:Residues: 19-267 <RES>  
 A:Cross-references: GB:M29068; NID:q178774; PIDN:AAA51747.1; PID:q178775  
 R:Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B.  
 J. Biol. Chem. 263, 18530-18536, 1988  
 A:Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by the APOA1 gene  
 A:Reference number: I39475; MUID:89054040  
 A:Accession: I39475  
 A:Molecule type: DNA  
 A:Residues: 1-14 <RE2>  
 A:Cross-references: GB:J04066; NID:q178763; PIDN:AAA51746.1; PID:q553183  
 R:Breslow, J.L.  
 Annu. Rev. Biochem. 54, 699-727, 1985  
 A:Title: Human apolipoprotein molecular biology and genetic variation.  
 A:Reference number: A90042; MUID:85278004  
 A:Contents: annotation; review of sequences, variants and gene location  
 R:Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B.  
 J. Biol. Chem. 261, 3911-3914, 1986  
 A:Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation  
 A:Reference number: A92577; MUID:86140194  
 A:Contents: annotation; acylation with palmitate  
 A:Note: an undetermined serine or threonine is acylated by fatty acid; the acylating agent is not determined  
 R:Law, S.W.; Brewer, H.B.  
 J. Biol. Chem. 260, 12810-12814, 1985  
 A:Title: Tangier disease: The complete mRNA sequence encoding for preproapo-A-I.  
 A:Reference number: I55236; MUID:86008382  
 A:Accession: I55236  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-143, 'D', '145-267 <RE3>  
 A:Cross-references: GB:M11791; NID:q178776; PIDN:AAA35545.1; PID:q178777  
 C:Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestine  
 A:Genetics: GDB:APOA1  
 A:Cross-references: GDB:119684; OMIM:107680  
 A:Map position: 11q23.3-11q23.3  
 A:Introns: 15/1; 67/2  
 C:Function: participates in the reverse transport of cholesterol from tissues to the liver  
 A:Description: participates in the reverse transport of cholesterol from tissues to the liver; noncovalently binds and stabilizes prostacyclin (PGI-2) and prostacyclin synthase (PGIS)

C;Superfamily: apolipoprotein A-I  
 C;Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipi  
 F;1-18/Domain: signal sequence #status experimental <SIG>  
 F;19-24/Domain: propeptide #status experimental <PRO>  
 F;25-267/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 100.0%; Score 468; DB 1; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-38;  
 Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKL 60  
 |||||  
 Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKL 84  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKD 89  
 |||||  
 Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKD 113

## RESULT 2

A26529

apolipoprotein A-I precursor - crab-eating macaque

C;Species: Macaca fascicularis (crab-eating macaque)

C;Date: 30-Sep-1989 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999

C;Accession: A26529; A26627; S23135; A57766

R;Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marotli, K.R.

Gene 49, 103-110, 1986

A;Title: The primary structure of cynomolgus monkey apolipoprotein A-I deduced from the

A;Reference number: A26529; MUID:87191989

A;Accession: A26529

A;Molecule type: mRNA

A;Residues: 1-267 &lt;POL&gt;

A;Cross-references: GB:M15411; NID:g342074; PIDN:AAA36834.1; PID:g342075

R;Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Sartelli, A.L.; Kantor, M.A.; Nicolosi,

Biochim. Biophys. Acta 1131, 207-210, 1992

A;Title: Nucleotide sequence of the cynomolgus monkey apolipoproteins in the Macaca fascicularis (cyn

A;Reference number: A26627; MUID:87185451

A;Accession: A26627

A;Molecule type: DNA

A;Residues: 1-12, 'L', 14-267 &lt;MUR&gt;

A;Cross-references: GB:M83242; NID:g342070; PIDN:AAA36832.1; PID:g342071

R;Sorci-Thomas, M.; Kearns, M.W.

J. Biol. Chem. 266, 18045-18050, 1991

A;Title: Transcriptional regulation of the apolipoprotein A-I gene.

A;Reference number: A57766; MUID:92011532

A;Accession: A57766

A;Molecule type: DNA

A;Residues: 1-10 &lt;RES&gt;

A;Cross-references: GB:M69223; NID:g342066; PIDN:AAA36831.1; PID:g553820

C;Comment: The precursor is synthesized in the liver and small intestine. The propeptide

F;1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F;19-24/Domain: propeptide #status predicted &lt;PPT&gt;

F;25-267/Product: apolipoprotein A-I #status predicted &lt;MAT&gt;

C;Genetics:

A;Introns: 15/1; 67/2

C;Superfamily: apolipoprotein A-I

C;Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe

F;1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F;19-24/Domain: propeptide #status predicted &lt;PPT&gt;

F;25-267/Product: apolipoprotein A-I #status predicted &lt;MAT&gt;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKL 60  
 |||||  
 Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKL 84  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKD 89  
 |||||  
 Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKD 113

## RESULT 3

JS0079

apolipoprotein A-I precursor - baboon

C;Species: Papio sp. (baboon)

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 13-Jun-1997

C;Accession: JS0079

R;Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; Vandenberg, J.L.

Gene 74, 483-490, 1988

A;Title: The baboon gene for apolipoprotein A-I: characterization of a cDNA clone and

A;Reference number: JS0079; MUID:89232739

A;Accession: JS0079

A;Molecule type: mRNA

A;Residues: 1-267 &lt;HIX&gt;

A;Experimental source: liver

C;Comment: This protein is the principal protein component of high density lipoprotei

C;Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase

C;Genetics: This protein contains a region of repeated amino acids which form amphipat

A;Gene: apoA1

C;Superfamily: apolipoprotein A-I

C;Keywords: HDL; lipid binding; lipoprotein

F;1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F;19-267/Product: apolipoprotein A-I #status predicted &lt;LAI&gt;

F;123-144, 145-166, 167-188, 189-210, 211-232, 233-254/Region: tandem repeats

## Query Match

Best Local Similarity 94.7%; Score 443; DB 2; Length 267;  
 Matches 83; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKL 60  
 |||||  
 Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKL 84  
 QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKD 89  
 |||||  
 Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKD 113

## RESULT 4

LPPGAI

apolipoprotein A-I precursor - dog

C;Species: Canis lupus familiaris (dog)

C;Date: 17-Dec-1982 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C;Accession: A60940; A03092; A61418

R;Luo, C.C.; Li, W.H.; Chan, L.

J. Lipid Res. 30, 1735-1746, 1989

A;Title: Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implic

A;Reference number: A60940; MUID:90132271

A;Accession: A60940

A;Molecule type: mRNA

A;Residues: 1-266 &lt;LUO&gt;

R;Chung, H.; Randolp, A.; Reardon, I.; Heinrichson, R.L.

J. Biol. Chem. 257, 2961-2967, 1982

A;Title: The covalent structure of apolipoprotein A-I from canine high density lipopr

A;Reference number: A03092; MUID:82142425

A;Accession: A03092

A;Molecule type: protein

A;Residues: 25-167, 'G', 169-201, 'Q', 203-234, 'Q', 236-266 &lt;CHU&gt;

R;Nakai, T.; Whayne, T.F.; Tang, J.

FEBS Lett. 64, 409-411, 1976

A;Title: The amino- and carboxyl-terminal sequences of canine apolipoprotein A-I.

A;Reference number: A61418; MUID:76210910

A;Accession: A61418





A:Title: Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit apolipoprotein A-I  
A:Reference number: S00230; MUID:88082866  
A:Accession: S00230  
A:Molecule type: mRNA  
A:Residues: 1-265 <PAN>  
A:Cross-references: EMBL:X06658; NID:g1461; PIDN:CAA29857.1; PID:g1462  
A:Note: the authors translated the codon AGC for residue 174 as Arg  
A:Accession: S20557  
A:Molecule type: DNA  
A:Residues: 1-17, 'R', 19-44, 'I', 46-122, 'Y', 124-146, 'V', 148-265 <PAN2>  
A:Cross-references: EMBL:X06659; NID:g1459; PIDN:CAA29858.1; PID:g1460  
C:Comment: This protein is synthesized in the small intestine.  
C:Comment: This protein is a major component of the high density lipoproteins in plasma.  
C:Genetics:  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repeat  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 81.8%; Score 383; DB 1; Length 265;  
Best Local Similarity 78.4%; Pred. No. 5.6e-30;  
Matches 69; Conservative 12; Mismatches 7; Indels 0; Gaps 0;  
QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDVSTFESKLR 61  
Db 25 DEPRSSWDKIKDFATVYVDVTKDSGREYVAQFEASAFGKQLNLKLDNWDVSLSTVSKLQ 84  
QY 62 EQLGPVTQEFWDNLEKETEGRLQEMSKD 89  
Db 85 EQLGPVTQEFWDNLEKETEGRLQEMSKD 112

## RESULT 9

A56858  
apolipoprotein A-I precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 05-Jan-1996 #sequence\_revision 23-Aug-1997 #text\_change 13-Aug-1999  
C:Accession: I45853; A56858; A34649  
R:O'Huigin, C.; Chan, L.; Li, W.  
Mol. Biol. Evol. 7, 327-339, 1990  
A:Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolution  
A:Reference number: I45853; MUID:90348478  
A:Accession: I45853  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-265 <OXH>  
A:Cross-references: GB:M35870; NID:g162677; PIDN:AAA30381.1; PID:g162678  
R:Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Auboirton, S.; Bauchart, D.; Chapman, M.J.; Gc  
Biochim. Biophys. Acta 1123, 145-150, 1992  
A:Title: Plasma lipid transport in the pre-ruminant calf, Bos spp: primary structure of h  
A:Reference number: A56858; MUID:92153895  
A:Accession: A56858  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 19-184, 'OL', 187-265 <SPA>  
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
A:Note: sequence extracted from NCBI backbone (NCBIP:83520)  
R:Auboirton, S.; Sparrow, D.A.; Beaubatie, L.; Bauchart, D.; Sparrow, J.T.; Laplaud, P.M.  
Biochem. Biophys. Res. Commun. 166, 833-839, 1990  
A:Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasma h  
A:Reference number: A34649; MUID:90147795  
A:Accession: A34649  
A:Molecule type: protein  
A:Residues: 25-70 <AUB>  
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
C:Superfamily: apolipoprotein A-I  
C:Keywords: lipid binding; lipoprotein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 81.8%; Score 383; DB 2; Length 265;  
Best Local Similarity 80.7%; Pred. No. 5.6e-30;  
Matches 71; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDVSTFESKLR 61  
Db 25 DDPOSSWDRVKDFATVYVEAIKDSGRDYVAQFEASALGKQLNLKLDNWDVTLASTLSKVR 84  
QY 62 EQLGPVTQEFWDNLEKETEGRLQEMSKD 89  
Db 85 EQLGPVTQEFWDNLEKETASTLRQEMHKD 112

## RESULT 10

LPRB12  
apolipoprotein A-I precursor (clone 22ap AI) - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S06064  
R:Paraskevopoulou, T.B.; Kritsis, A.; Zannis, V.  
submitted to the EMBL Data Library, July 1989  
A:Reference number: S06064  
A:Accession: S06064  
A:Molecule type: mRNA  
A:Residues: 1-266 <PAR>  
A:Cross-references: EMBL:X15908; NID:g1457; PIDN:CAA34024.1; PID:g1458  
C:Comment: This protein is synthesized in the small intestine.  
C:Comment: This protein is a major component of the high density lipoproteins in plas  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 81.8%; Score 383; DB 1; Length 266;  
Best Local Similarity 78.4%; Pred. No. 5.6e-30;  
Matches 69; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDVSTFESKLR 61  
Db 25 DEPRSSWDKIKDFATVYVDVTKDSGREYVAQFEASAFGKQLNLKLDNWDVSLSTVSKLQ 84  
QY 62 EQLGPVTQEFWDNLEKETEGRLQEMSKD 89  
Db 85 EQLGPVTQEFWDNLEKETEGRLQEMSKD 112

## RESULT 11

A24998  
apolipoprotein A-I - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 31-Dec-1993  
C:Accession: A24998  
R:Yang, C.; Yang, T.; Pownall, H.J.; Gotto Jr., A.M.  
Eur. J. Biochem. 160, 427-431, 1986  
A:Title: The primary structure of apolipoprotein A-I from rabbit high-density lipopro  
A:Reference number: A24998; MUID:87030294  
A:Accession: A24998  
A:Molecule type: protein  
A:Residues: 1-241 <YAN>  
C:Superfamily: apolipoprotein A-I  
C:Keywords: HDL; lipid binding; lipoprotein

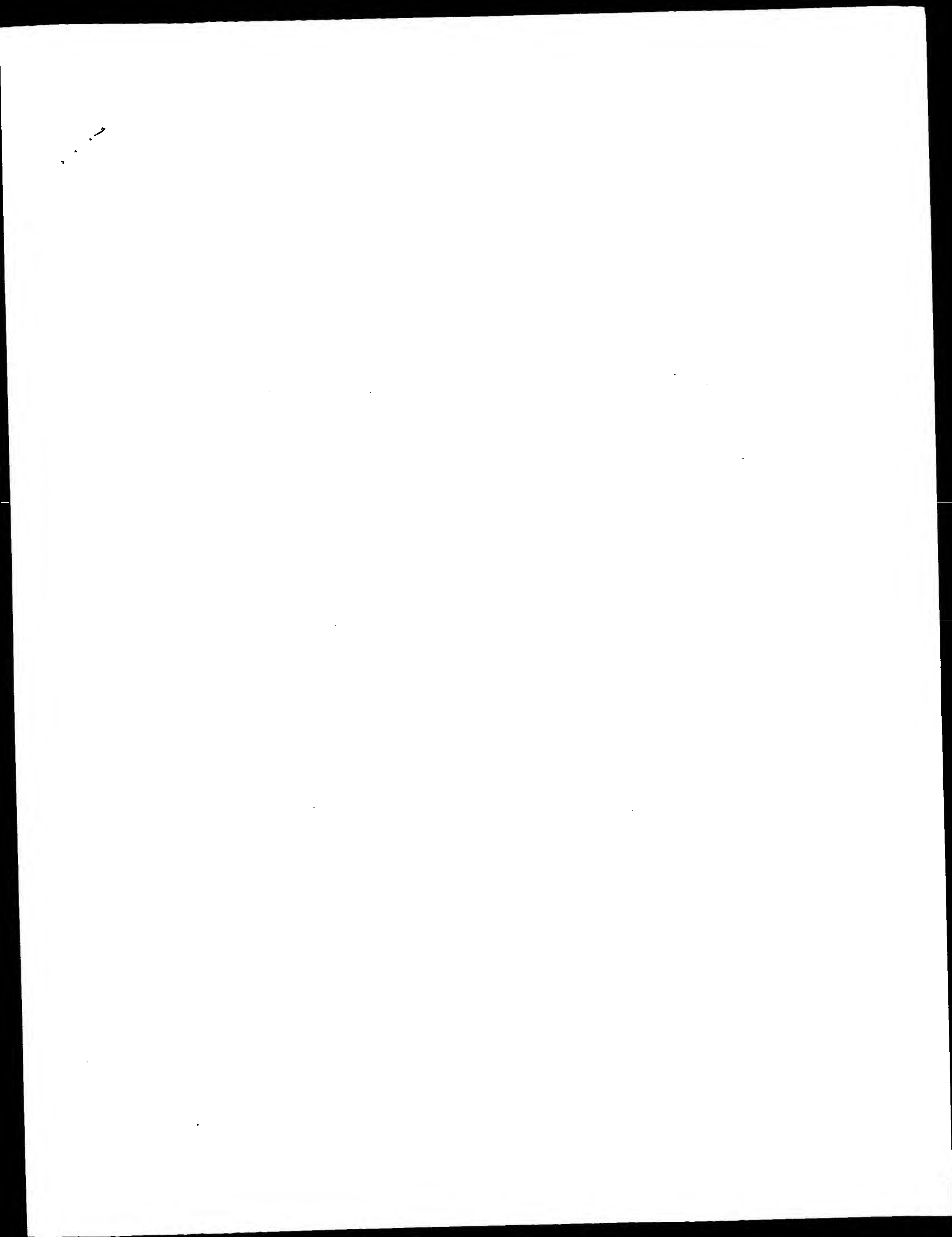
Query Match 80.9%; Score 378.5; DB 2; Length 241;  
Best Local Similarity 80.7%; Pred. No. 1.4e-29;  
Matches 71; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDVSTFESKLR 61  
Db 1 DEPRSSWDKIKDFATVYVDV-KDSGREYVAQFEASAFGKQLNLKLDNWDVSLSTVSKLQ 59









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:22:18 ; Search time 35.02 Seconds

(without alignments)  
98.402 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_113

Perfect score: 468  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....EFWDNLEKETEGLRQEMSKD 89

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	267	1	APAI_HUMAN P02647 homo sapien
2	443	94.7	267	1	APAI_MACFA P15568 macaca fasc
3	407	87.0	266	1	APAI_CANFA P02648 canis fam11
4	405	86.5	265	1	P18648 sus scrofa
5	383	81.8	265	1	APAI_BOVIN P15497 bos taurus
6	383	81.8	266	1	APAI_RABIT P09809 oryctolagus
7	359	76.7	259	1	APAI_RAT P04639 rattus norv
8	343	73.3	264	1	APAI_MOUSE Q00623 mus musculu
9	341	72.9	265	1	APAI_TUPGB O18759 tupala glis
10	223	47.6	264	1	APAI_CHICK P08250 gallus gall
11	223	47.6	264	1	APAI_COTJA P32918 coturnix co
12	220	47.0	264	1	O42296 anas platyr
13	101	21.6	20	1	APAI_ERIPA P18647 erythroceu
14	92	19.7	396	1	APAI_HUMAN P06727 homo sapien
15	90	19.2	391	1	APAI_RAT P02651 rattus norv
16	90	19.2	429	1	APAI_PAPAN P33621 macaca fasc
17	86	18.4	401	1	APAI_BRARE Q28758 papio anubi
18	84	17.9	281	1	O42364 brachydanio
19	83	17.7	312	1	P02650 rattus norv
20	82	17.5	382	1	APAI_PIG O46409 sus scrofa
21	81	17.3	365	1	RECF_CHLTR O84077 chlamydia t
22	81	17.3	395	1	APAI_MOUSE P06728 mus musculu
23	78	16.7	311	1	APAI_MOUSE P08226 mus musculu
24	76.5	16.3	207	1	YEE6_YEAST P43557 saccharomyc
25	72.5	15.5	192	1	YE49_ANASP P29979 anabaena sp
26	70.5	15.1	707	1	HS88_NEUCR O74225 neurospora
27	70	15.0	452	1	TNA2_SYMTM P31015 symbiobacte
28	70	15.0	453	1	TNA1_SYMTM P31014 symbiobacte
29	69.5	14.9	759	1	GIT2_HUMAN O14161 homo sapien
30	69	14.7	2033	1	EVPL_HUMAN Q92817 homo sapien
31	68.5	14.6	1453	1	MLH3_HUMAN Q9uhc1 homo sapien
32	68.5	14.6	4725	1	DYHC_DICDI P34036 dictyosteli
33	67.5	14.4	686	1	MEPD_RAT P24155 rattus norv

34	66.5	14.2	185	1	GERD_BACSU P16450 bacillus su
35	66.5	14.2	459	1	FV1_MOUSE P70213 mus musculu
36	66	14.1	210	1	YG21_YEAST P53251 saccharomyc
37	66	14.1	298	1	APAI_CAVPO P23529 cavia porce
38	66	14.1	317	1	APAI_PIG P18650 sus scrofa
39	66	14.1	365	1	RECF_CHLMU Q9pkw5 chlamydia m
40	66	14.1	1051	1	CARB_SULSO Q59969 sulfolobus
41	65.5	14.0	261	1	HB23_HUMAN P05537 homo sapien
42	65.5	14.0	550	1	G6PI_SCHPO P78917 schizosacch
43	65	13.9	203	1	GTSL_BLAGO O18598 blatella g
44	65	13.9	294	1	CPEM_STRHY P11435 streptomyc
45	64.5	13.8	747	1	AT12_HSVB P28937 equine herp

## ALIGNMENTS

RESULT	ID	APAI_HUMAN	STANDARD;	PRT;	267 AA.
AC	P02647;				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).				
GN	APOA1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
[11]					
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85026665; PubMed=6207999;				
RA	Seilhamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;				
RT	"Isolation and DNA sequence of full-length cDNA and of the entire				
RT	gene for human apolipoprotein AI -- discovery of a new genetic				
RT	polymorphism in the apo AI gene.";				
RL	DNA 3:309-317(1984).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=83220822; PubMed=6406984;				
RA	Shoulders C.C., Kornblith A.R., Munro B.S., Baralle F.E.;				
RT	"Gene structure of human apolipoprotein AI.";				
RL	Nucleic Acids Res. 11:2827-2837(1983).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=83220772; PubMed=6304641;				
RA	Cheung P., Chan L.;				
RT	"Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";				
RL	Nucleic Acids Res. 11:3703-3715(1983).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84119464; PubMed=6198645;				
RA	Law S.W., Brewer H.B. Jr.;				
RT	"Nucleotide sequence and the encoded amino acids of human				
RT	apolipoprotein A-I mRNA.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).				
RN	[6]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=86008382; PubMed=2995392;				
RA	Law S.W., Brewer H.B. Jr.;				
RT	"Tangier disease. The complete mRNA sequence encoding for				
RT	preproapo-A-I.";				
RL	J. Biol. Chem. 260:12810-12814(1985).				
RN	[7]				



RP SEQUENCE FROM N.A.  
 RX MEDLINE=84016011; PubMed=6413973;  
 RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
 RT "Isolation and characterization of the human apolipoprotein A-I  
 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89377481; PubMed=2673706;  
 RA Mogilevsky N., Roobol C., Lorian R., Guillaume J.P., Jacobs P.,  
 RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
 RA Holmquist L., Carlson L.A., Bollen A.;  
 RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
 coli: purification and biochemical characterization.";  
 RL DNA 8:429-436(1989).  
 RN [9]  
 RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
 RX MEDLINE=88196137; PubMed=3129297;  
 RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
 RA Zannis V.I.;  
 RT "Sequence and expression of Tangier apoA-I gene.";  
 RL Eur. J. Biochem. 173:465-471(1988).  
 RN [10]  
 RP SEQUENCE OF 118-267 FROM N.A.  
 RX MEDLINE=83091059; PubMed=6294659;  
 RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
 RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
 RT "Isolation and characterization of cDNA clones for human  
 apolipoprotein A-I.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
 RN [11]  
 RP SEQUENCE OF 19-27.  
 RX MEDLINE=83256553; PubMed=6409108;  
 RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
 RA Light J.A.;  
 RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
 RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
 RN [12]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=78123731; PubMed=204308;  
 RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
 RA Bronzert T.J.;  
 RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated  
 from high density lipoproteins.";  
 RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
 RN [13]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=75133493; PubMed=164450;  
 RA Baker H.N., Gatto A.M. Jr., Jackson R.L.;  
 RT "The primary structure of human plasma high density apolipoprotein  
 glutamine I (ApoA-I). II. The amino acid sequence and alignment of  
 cyanogen bromide fragments IV, III, and I.";  
 RL J. Biol. Chem. 250:2725-2738(1975).  
 RN [14]  
 RP SEQUENCE OF 25-56.  
 RX MEDLINE=88331387; PubMed=3047170;  
 RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
 RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
 A-I (Apo A-I). A novel function of Apo A-I.";  
 RL J. Clin. Invest. 82:803-807(1988).  
 RN [15]  
 RP SEQUENCE OF 25-48.  
 RX MEDLINE=89380318; PubMed=2506184;  
 RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
 RA Chadelaine A.;  
 RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
 proteins.";  
 RL J. Biol. Chem. 264:16853-16857(1989).  
 RN [16]  
 RP SEQUENCE OF 25-43.  
 RX MEDLINE=88070603; PubMed=3120314;  
 RA Priolo R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
 RA Pereira M.E.A.;

RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
 RT neuraminidase, to high-density lipoprotein.";  
 RL Science 238:1417-1419(1987).  
 RN [17]  
 RP SEQUENCE OF 25-42.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 RN [18]  
 RP PALMITOYLATION.  
 RX MEDLINE=86140194; PubMed=3005308;  
 RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
 RT "Human apolipoprotein A-I. Post-translational modification by fatty  
 RT acid acylation.";  
 RL J. Biol. Chem. 261:3911-3914(1986).  
 RN [19]  
 RP PROCESSING.  
 RX MEDLINE=83195100; PubMed=6405383;  
 RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
 RA Breslow J.L.;  
 RT "Intracellular and extracellular processing of human apolipoprotein  
 RT A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
 RN [20]  
 RP STRUCTURE BY NMR OF 190-209.  
 RX MEDLINE=96270776; PubMed=8664326;  
 RA Wang G., Treleaven W.D., Cushley R.J.;  
 RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
 RT presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
 RT and CD. Evidence for specific peptide-SDS interactions.";  
 RL Biochim. Biophys. Acta 1301:174-184(1996).  
 RN [21]  
 RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
 RX MEDLINE=98024124; PubMed=9356442;  
 RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
 RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
 RT lipid-bound conformation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
 RN [22]  
 RP VARIANT MILANO.  
 RX MEDLINE=83109095; PubMed=6401735;  
 RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,  
 RA Franceschini G., Sirtori C.R.;  
 RT "Apolipoprotein A-IMilano. Detection of normal A-I in affected  
 RT subjects and evidence for a cysteine for arginine substitution in the  
 RT variant A-I.";  
 RL J. Biol. Chem. 258:2508-2513(1983).  
 RN [23]  
 RP VARIANT TANGIER.  
 RX MEDLINE=83300108; PubMed=6412234;  
 RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;  
 RT "Tangier disease: defective recombination of a specific Tangier  
 RT apolipoprotein A-I isoform (pro-apo A-I) with high density  
 RT lipoproteins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).  
 RN [24]  
 RP VARIANT NORWAY.  
 RX MEDLINE=84289383; PubMed=6432779;  
 RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,  
 RA Utermann G., Haas J., Steinmetz A., Menzel H.J., Assmann G.;  
 RT "Abnormal lecithin:cholesterol acyltransferase activation by a human  
 RT apolipoprotein A-I variant in which a single lysine residue is  
 RT deleted.";  
 RL J. Biol. Chem. 259:10063-10070(1984).  
 RN [25]  
 RP SEQUENCE OF 25-107 (VARIANT IOWA).  
 RX MEDLINE=89050104; PubMed=3142462;  
 RA Nichols W.C., Dwulet F.E., Liepnieks J., Benson M.D.;  
 RT "Variant apolipoprotein AI as a major constituent of a human  
 RT hereditary amyloid.";

```
Query Match          100.0%; Score 468; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.9e-39;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSL 60
    |||
Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSL 84
QY 61 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 89
    |||
Db 85 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 113

RESULT 2
APAL_MACFA
ID APAL_MACFA STANDARD; PRT; 267 AA.
AC P15568; P17929;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541, 9557;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.fascicularis;
RX MEDLINE=87191989; PubMed=3106152;
RA Polites H.G., Melchior G.W., Castle C.K., Marotti K.R.;
RT "The primary structure of cynomolgus monkey apolipoprotein A-1
RL deduced from the cDNA sequence: comparison to the human sequence.";
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.fascicularis;
RX MEDLINE=92305062; PubMed=1610902;
RA Murray R.W., Marotti K.R.;
RT "Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene
RL and corresponding flanking regions.";
RN [3]
RP SEQUENCE OF 25-48.
RC SPECIES=M.fascicularis;
RX MEDLINE=87185451; PubMed=3105581;
RA Herbert P.N., Bausserman L.L., Lynch K.M., Saritelli A.L.,
RA Kantor M.A., Nicolosi R.J., Shulman R.S.;
RT "Homologues of the human C and A apolipoproteins in the Macaca
RL fascicularis (cynomolgus) monkey.";
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=P.hamadryas; TISSUE=Liver;
RX MEDLINE=89232739; PubMed=2907746;
RA Hixson J.E., Borenstein S., Cox L.A., Rainwater D.L., Vandeberg J.L.;
RT "The baboon gene for apolipoprotein A-I: characterization of a cDNA
RT clone and identification of DNA polymorphisms for genetic studies of
RT cholesterol metabolism.";
RN [5]
RP SEQUENCE OF 1-10 FROM N.A.
RC SPECIES=M.fascicularis;
RA Sorci-Thomas M.;
RL Submitted (OCT-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
```

```
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15411; AAA36834.1; -
DR EMBL; M83242; AAA36832.1; -
DR EMBL; M35634; AAA35380.1; -
DR EMBL; M69223; AAA36831.1; -
DR PIR; A26529; A26529.
DR PIR; A26627; A26627.
DR PIR; JS0079; JS0079.
DR PIR; S23135; S23135.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 267
FT DOMAIN 68 267
FT REPEAT 68 89
FT REPEAT 90 111
FT REPEAT 112 122
FT REPEAT 123 144
FT REPEAT 145 166
FT REPEAT 167 188
FT REPEAT 189 210
FT REPEAT 211 232
FT REPEAT 233 243
FT REPEAT 244 267
FT CONFLICT 13 13
FT SEQUENCE 267 AA; 30735 MW; 869955C024088E21 CRC64;

Query Match          94.7%; Score 443; DB 1; Length 267;
Best Local Similarity 93.3%; Pred. No. 1.1e-36;
Matches 83; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSL 60
    |||
Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSL 84
QY 61 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 89
    |||
Db 85 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 113

RESULT 3
APAL_CANFA
ID APAL_CANFA STANDARD; PRT; 266 AA.
AC P02648;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90132271; PubMed=2515239;
RA Luo C.-C., Li W.-H., Chan L.;
```

RT "Structure and expression of dog apolipoprotein A-I, E, and C-I  
RT mRNAs: implications for the evolution and functional constraints of  
RT apolipoprotein structure.";  
RL J. Lipid Res. 30:1735-1746(1989).  
RN [2]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=82142425; PubMed=6801039;  
RA Chung H., Randalph A., Reardon I., Heinrichson R.L.;  
RT "The covalent structure of apolipoprotein A-I from canine high  
RT density lipoproteins.";  
RL J. Biol. Chem. 257:2961-2967(1982).  
RN [3]  
RP SEQUENCE OF 25-57 AND 262-265.  
RX MEDLINE=76210910; PubMed=179887;  
RA Nakai T., Whayne T.F., Tang J.;  
RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein  
RT A-I.";  
RL FEBS Lett. 64:409-411(1976).  
RN [4]  
RP SEQUENCE OF 25-37.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
RT dog heart proteins.";  
RL Electrophoresis 18:2795-2802(1997).  
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -I- SUBCELLULAR LOCATION: Extracellular.  
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS. SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.  
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
DR PIR: A03092; LPDGA1.  
DR HSSP: P02647; 1ODR.  
DR HSC-2DPAGE; P02648; DOG.  
DR InterPro; IPR00074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18 BY SIMILARITY.  
FT PROPEP 19 24  
FT CHAIN 25 266 APOLIPOPROTEIN A-I.  
FT DOMAIN 67 266 10 X APPROXIMATE TANDEM REPEATS.  
FT REPEAT 67 88 1.  
FT REPEAT 89 110 2.  
FT REPEAT 111 121 3 (HALF-LENGTH).  
FT REPEAT 122 143 4.  
FT REPEAT 144 165 5.  
FT REPEAT 166 187 6.  
FT REPEAT 188 209 7.  
FT REPEAT 210 231 8.  
FT REPEAT 232 242 9 (HALF-LENGTH).  
FT REPEAT 243 266 10.  
FT CONFLICT 168 168 A -> G (IN REF. 2).  
FT CONFLICT 202 202 E -> Q (IN REF. 2).  
FT CONFLICT 235 235 E -> Q (IN REF. 2).  
FT CONFLICT 264 266 NAQ -> A (IN REF. 3).  
SQ SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;

Query Match 87.0%; Score 407; DB 1; Length 266;  
Best Local Similarity 86.4%; Pred. No. 4e-33;  
Matches 76; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 EPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTFSKLR 61  
Db : |||||  
QY 25 DEPQSPMDRVKDLATVYVDVAVKDSGRDYVAFASALGKQLNLKLDNWDVSLSTVTKLR 84  
Db : |||||  
QY 62 EQLGPVTOEFWDNLEKETEGLRQEMSKD 89  
Db : |||||  
Db 85 EQIGPVTQEFWDNLEKETEVLRQEMSKD 112

RESULT 4  
ID APAL\_PIG STANDARD; PRT; 265 AA.  
AC P18648;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93224154; PubMed=8468059;  
RA Birchbauer A., Knipping G., Juritsch B., Aschauer H., Zechner R.;  
RT "Characterization of the apolipoprotein AI and CIII genes in the  
RT domestic pig.";  
RL Genomics 15:643-652(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Brain;  
RX MEDLINE=94125128; PubMed=8294940;  
RA Moeckel B., Zinke H., Flach R., Weiss B., Weller-Guettler H.,  
RT Gassen H.;  
RT "Expression of apolipoprotein A-I in porcine brain endothelium in  
RT vitro.";  
RL J. Neurochem. 62:788-798(1994).  
RN [3]  
RP SEQUENCE OF 34-265 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90132667; PubMed=2105375;  
RA Weller-Guettler H., Sommerfeldt M., Papandrikopoulou A., Mischek U.,  
RA Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;  
RT "Synthesis of apolipoprotein A-I in pig brain microvascular  
RT endothelial cells.";  
RL J. Neurochem. 54:444-450(1990).  
RN [4]  
RP SEQUENCE OF 105-265 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93154581; PubMed=8428656;  
RA Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;  
RT "Sequences and expression of the porcine apolipoprotein A-I and C-III  
RT mRNAs.";  
RL Gene 123:173-179(1993).  
RN [5]  
RP SEQUENCE OF 25-265.  
RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,  
RA Jackson K., Gustavsson I., Rapacz J.;  
RL Submitted (OCT-1995) to the SWISS-PROT data bank.  
RN [6]  
RP SEQUENCE OF 25-34.  
RX MEDLINE=76184721; PubMed=178359;  
RA Mahley R.W., Welsgraber K.H., Innerarity T., Brewer H.B. Jr.;  
RT "Characterization of the plasma lipoproteins and apoproteins of the  
RT Erythrocebus patas monkey.";  
RL Biochemistry 15:1928-1933(1976).  
CC -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -I- SUBCELLULAR LOCATION: Extracellular.  
CC -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE  
CC LIVER.  
CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way





RESULT	6			
APAL_RABIT				
ID	APAL_RABIT	STANDARD;	PRT;	266 AA.
AC	P09809;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Apolipoprotein A-I precursor (Apo-AI).			
GN	APoAI.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=2ZAP AI; TISSUE=Small intestine;			
RA	Paraskevopoulou T.B., Kritis A., Zannis V.I.;			
RL	Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Intestine;			
RX	MEDLINE=88082866; PubMed=3121329;			
RA	Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,			
RA	Kroon P.A., Chao Y.S.;			
RT	"Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit			
RT	apolipoprotein A-I is synthesized in the intestine but not in the			
RT	liver.";			
RL	Eur. J. Biochem. 170:99-104(1987).			
RN	[3]			
RP	SEQUENCE OF 25-266.			
RX	MEDLINE=87030294; PubMed=3095115;			
RA	Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;			
RT	"The primary structure of apolipoprotein A-I from rabbit high-density			
RT	lipoprotein.";			
RL	Eur. J. Biochem. 160:427-431(1986).			
CC	-I- FUNCTION: APOA-I PARTICIPATES IN THE REVERSE TRANSPORT OF			
CC	CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING			
CC	CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR			
CC	THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).			
CC	-I- SUBCELLULAR LOCATION: Extracellular.			
CC	-I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN			
CC	CHYLOMICRONS.			
CC	-I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL cutstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X15908; CAA34024.1; -			
DR	EMBL; X06658; CAA29857.1; -			
DR	EMBL; X06659; CAA29858.1; -			
DR	PIR; S06064; LPRBIZ.			
DR	PIR; S00230; LPRBIB.			
DR	PIR; A24998; A24998.			
DR	HSSP; P02647; IAV1.			
DR	InterPro; IPR000074; Apolipoprotein.			
DR	pfam; PF01442; Apolipoprotein; 1.			
KW	Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.			
FT	SIGNAL	1	18	
FT	PROPEP	19	24	
FT	CHAIN	25	266	APOLIPOPROTEIN A-I.
FT	DOMAIN	67	266	10 X APPROXIMATE TANDEM REPEATS.
FT	REPEAT	67	88	1.
FT	REPEAT	89	110	2.
FT	REPEAT	111	121	3 (HALF-LENGTH).
FT	REPEAT	122	143	4.
FT	REPEAT	144	165	5.
FT	REPEAT	166	187	6.
FT	REPEAT	188	209	7.

FT	REPEAT	210	231		8.	(HALF-LENGTH).	
FT	REPEAT	232	242		9.		
FT	REPEAT	243	266		10.	A -> R (IN REF. 2; CAA29858).	
FT	CONFLICT	18	18			MISSING (IN REF. 3).	
FT	CONFLICT	44	44		V -> I (IN REF. 2; CAA29858).		
FT	CONFLICT	45	45		E -> Q (IN REF. 3).		
FT	CONFLICT	107	107		Y -> F (IN REF. 2; CAA29857).		
FT	CONFLICT	123	123		A -> V (IN REF. 2; CAA29858 AND REF. 3).		
FT	CONFLICT	147	147		R -> G (IN REF. 2; CAA29858 AND CAA29857).		
FT	CONFLICT	150	150		N -> Q (IN REF. 3).		
FT	CONFLICT	191	191		MISSING (IN REF. 2; CAA29858 AND CAA29857).		
FT	CONFLICT	195	195		S -> K (IN REF. 2; CAA29858 AND CAA29857).		
FT	CONFLICT	211	211		VL -> LV (IN REF. 3).		
FT	CONFLICT	255	256		L -> V (IN REF. 2; CAA29858 AND CAA29857).		
FT	CONFLICT	256	256				
SO	SEQUENCE	266 AA;	30591 MW;	0FF6DB386497C7D2 CRC64;			
QY	Query Match	Best Local Similarity	81.8%; Score 383; DB 1; Length 266;				
Db	Matches	69; Conservative	12; Mismatches	7; Indels	0; Gaps	0;	
QY	2 EPPSPWDRVKDLATVYVDYLKDSGRDYVSQFEGSALGQLNLKLNDWSVTSTFSKL	61	: ::            ::   ::                      :				
Db	25 DEPRSSWDKIKDFATVVYDVTKDSGREYVAQFEASAFGKQLNKLNDWSLSTSVSKLQ	84	:  :				
QY	62 EQLGPTVOEFWDNLKETEGLRQEMSKD	89	:				
Db	85 EOLGPVTQOEFWDNLKETEGLRQEMNKD	112					
RESULT	7						
APAI_RAT	STANDARD;	PRT;	259 AA.				
ID APAI_RAT	P04639;						
AC	13-AUG-1987 (Rel. 05, Created)						
DT	13-AUG-1987 (Rel. 05, Last sequence update)						
DT	30-MAY-2000 (Rel. 39, Last annotation update)						
DE	Apolipoprotein A-I precursor (Apo-AI).						
GOS	APoAI.						
OS	Rattus norvegicus (Rat).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.						
OX	NCBI_TaxID=10116;						
RP	[1]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=84207987; PubMed=6426956;						
RA	Poncin J.E., Martial J.A., Gielen J.E.;						
RL	"Cloning and structure analysis of the rat apolipoprotein A-I cDNA.";						
RN	Eur. J. Biochem. 140:493-498(1984).						
RP	[2]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=87008540; Pubmed=3020028;						
RA	Haddad I.A., Ordovas J.M., Fitzpatrick T., Karathanasis S.K.;						
RT	"Linkage, evolution, and expression of the rat apolipoprotein A-I, C-						
RT	III, and A-IV genes.";						
RL	J. Biol. Chem. 261:13268-13277(1986).						
RN	[3]						
RP	SEQUENCE OF 1-45.						
RC	STRAIN=SPRAGUE-DAWLEY;						
RX	MEDLINE=82098162; Pubmed=6798036;						
RA	Gordon J.I., Smith D.P., Andy R., Alpers D.H., Schonfeld G.,						
RA	Strauss A.W.;						
RT	"The primary translation product of rat intestinal apolipoprotein A-I						
RT	mRNA is an unusual preproprotein.";						
CC	J. Biol. Chem. 257:971-978(1982).						
CC	-! FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF						
CC	CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING						
CC	CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR						

```

CC      THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC      -1- SUBCELLULAR LOCATION: Extracellular.
CC      -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC      CHYLOMICRONS.
CC      -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M00001; AAA40749.1; -.
DR      EMBL; X00558; CAA25224.1; -.
DR      EMBL; J02597; AAA40745.1; -.
DR      PIR; A24700; A24700.
DR      HSSP; P02647; IAVI.
DR      InterPro; IPR00074; Apolipoprotein.
DR      Pfam; PF01442; Apolipoprotein; 1.
KW      Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT      SIGNAL          1      18
FT      PROPEP          19      24
FT      CHAIN           25      259
FT      DOMAIN          67      259
FT      REPEAT          67      88
FT      REPEAT          89      110
FT      REPEAT          111     121
FT      REPEAT          122     143
FT      REPEAT          144     161
FT      REPEAT          162     183
FT      REPEAT          184     203
FT      REPEAT          204     225
FT      REPEAT          226     236
FT      REPEAT          237     259
FT      REPEAT          259     201
FT      CONFLICT        201     214
FT      CONFLICT        214     218
FT      CONFLICT        218     218
SQ      SEQUENCE        259 AA; 30088 MW; 2E8D5EB45FEAE88 CRC64;

Query Match          76.7%; Score 359; DB 1; Length 259;
Best Local Similarity 76.1%; Pred. No. 2e-28;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY      2 EPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGQNLKLLDNWDSVTSFSKLR 61
      : ||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      25 DEPOSQWDRVKDEFAIYVDVAVKDSGRDYSQFESSTLGKQLNLNLNDWDTLGGTVGRLQ 84

QY      62 EQLGPTQEFWDNLEKTEGLRQEMSKD 89
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      85 EQLGPTQEFWANLEKETDWLRNEMNKD 112

RESULT 8
APAI_MOUSE
ID      APAI_MOUSE          STANDARD;          PRT;          264 AA.
AC      Q00623;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Apolipoprotein A-I precursor (Apo-AI).
GN      APOA1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92281682; PubMed=1596360;
RA      Stoffel W., Mueller R., Binczek E., Hofmann K.;
      *Mouse apolipoprotein AI. cDNA-derived primary structure, gene

```

```

RT      organisation and complete nucleotide sequence.";
RL      Biol. Chem. Hoppe-Seyler 373:187-193(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93122774; PubMed=1478650;
RA      Januzzi J.L., Azrolan N., O'Connell A., Aalto-Setälä K., Breslow J.L.;
RT      "Characterization of the mouse apolipoprotein ApoA-1/ApoC-3 gene
RT      locus: genomic, mRNA, and protein sequences with comparisons to other
RT      species.";
RL      Genomics 14:1081-1088(1992).
CC      -I- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC      CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC      CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC      THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC      -I- SUBCELLULAR LOCATION: Extracellular.
CC      -I- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC      CHYLOMICRONS.
CC      -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X64262; CAA45560.1; -.
DR      EMBL; X64263; CAA45561.1; -.
DR      EMBL; L04149; -; NOT_ANNOTATED_CDS.
DR      EMBL; L04151; -; NOT_ANNOTATED_CDS.
DR      PIR; S22420; S22420.
DR      PIR; A44364; A44364.
DR      HSSP; P02647; LAVI.
DR      SWISS-2DPAGE; Q00623; MOUSE.
DR      MGD; MGI:88049; Apoa1.
DR      InterPro; IPR000074; Apolipoprotein.
DR      Pfam; PF01442; Apolipoprotein; 1.
KW      Plasma, Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT      SIGNAL          1           18       BY SIMILARITY.
FT      PROPEP         19           24       BY SIMILARITY.
FT      CHAIN          25           264      APOLIPOPROTEIN A-I.
FT      DOMAIN        67           264      10 X APPROXIMATE TANDEM REPEATS.
FT      REPEAT         67            88       1.
FT      REPEAT         89           110       2.
FT      REPEAT        111           121       3 (HALF-LENGTH).
FT      REPEAT        122           143       4.
FT      REPEAT        144           165       5.
FT      REPEAT        166           187       6.
FT      REPEAT        188           207       7 (INCOMPLETE).
FT      REPEAT        208           229       8.
FT      REPEAT        230           240       9 (HALF-LENGTH).
FT      REPEAT        241           264      10.
SQ      SEQUENCE      264 AA; 30587 MW; C453FF2019634AAC CRC64;

Query Match          73.3%; Score 343; DB 1; Length 264;
Best Local Similarity 69.3%; Pred. No. 7.8e-27;
Matches 61; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY      2 EPQPSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSVTSFSKL R 61
DB      : ||| |||| | ||| :||||||| | |:||| | | ||| : ||| :
DB      25 DEPOSQMDKVDFANVYDVAVKDSGRDYVSQFESSSLGQQLNLLENWDTLGSIYSQIQ 84
QY      62 EQLGPVTQEFWDNLEKETEGLRQEMSKD 89
DB      !:|||!::|!!!!!!|: :!!!!!!|
DB      85 ERLGPLTRDFWNLEKETDVRQEMNKD 112

RESULT   9
APAI_TUPGB STANDARD; PRT; 265 AA.
AC 018759;
```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Tupaiia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;
RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF005638; AAB82326.1; -.
DR HSSP; P02647; 1A.V1.
DR InterPro; IPR00074; Apolipoprotein.
DR pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 265 APOLOPOPROTEIN A-I.
FT DOMAIN 67 265 10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (HALF-LENGTH).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (HALF-LENGTH).
FT REPEAT 243 265 10.
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAEA165 CRC64;

```

```

Query Match 72.9%; Score 341; DB 1; Length 265;
Best Local Similarity 70.5%; Pred. No. 1.2e-26;
Matches 62; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 2 EPPQSPWDRVKDIATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTVSTFSKLR 61
   : ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| |
Db 25 DEPOSSWDRVRLANVYVDVAVKESGREYVSQLLEASALGKQLNLKLDNWDSTLGSTFGKVH 84
QY 62 EQLGPVTQEFWNLKETEGLRQEMSKD 89
   | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 85 EHLGPVAQEFWEKLEKETELRREINKD 112

```

```

RESULT 10
APAI_CHICK STANDARD; PRT; 264 AA.
AC P08250;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88049703; PubMed=3118875;
RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.;
RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and
RT evolution.";
RL Biochem. Biophys. Res. Commun. 148:485-492(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88152500; PubMed=3126099;
RA Ferrari S., Tarugi P., Drusiani E., Calandra S., Fregni M.;
RT "The complete sequence of chick apolipoprotein AI mRNA and its
RT expression in the developing chick.";
RL Gene 60:39-46(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87222301; PubMed=3108248;
RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,
RA Leberer H., Lusis A.J.;
RT "Structure, evolution, and regulation of chicken apolipoprotein A-I.";
RL J. Biol. Chem. 262:7058-7065(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92381402; PubMed=1512510;
RA Lamou-Fava S., Sastri R., Ferrari S., Rajavashisth T.B.,
RA Lusis A.J., Karathanasis S.K.;
RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene
RT expression: differences between avian and mammalian apoA-I gene
RT transcription control regions.";
RL J. Lipid Res. 33:831-842(1992).
RN [5]
RP SEQUENCE OF 25-44.
RX MEDLINE=83213468; PubMed=6406496;
RA Shackelford J.E., Leberer H.G.;
RT "Synthesis and secretion of apolipoprotein AI by chick breast
RT muscle.";
RL J. Biol. Chem. 258:7175-7180(1983).
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
CC CHYLOMICRONS.
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17961; AAA48593.1; -.
DR EMBL; M18746; AAA48594.1; -.
DR EMBL; M25559; AAA48592.1; -.
DR EMBL; M96012; AAA48597.1; -.
DR PIR; S01453; LPCHA1.
DR PIR; JH0471; JH0471.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR00074; Apolipoprotein.
DR pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24

```







RA Lohse F., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to  
 RT nucleotide substitutions in the apolipoprotein A-IV gene.";  
 RL J. Biol. Chem. 265:10061-10064(1990).  
 RN [9]  
 RP VARIANTS A-IV\*0 AND A-IV\*3.  
 RX MEDLINE=90324273; PubMed=1973689;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for  
 RT two rare variants of apolipoprotein A-IV-1.";  
 RL J. Biol. Chem. 265:12734-12739(1990).  
 RN [10]  
 RP VARIANTS.  
 RX MEDLINE=91310615; PubMed=1677358;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV  
 RT 1(7th-347-->Ser), apoA-IV-0(Lys-167-->Glu,Gln-360-->His), and apoA-I  
 RT 3(Glu-165-->Lys).";  
 RL J. Biol. Chem. 266:13513-13518(1991).  
 RN [11]  
 RP ERRATUM.  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RL J. Biol. Chem. 266:19866-19866(1991).  
 RN [12]  
 RP VARIANT MET-13.  
 RX MEDLINE=92238494; PubMed=1349197;  
 RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
 RA Assmann G.;  
 RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
 RT gene are associated with changes in the concentration of apo B- and  
 RT apo A-I-containing lipoproteins in a normal population.";  
 RL Am. J. Hum. Genet. 50:1115-1128(1992).  
 RN [13]  
 RP VARIANT SER-147.  
 RX MEDLINE=92144647; PubMed=1737067;  
 RA Tenkanen H., Koskinen P., Metsä J., Baumann M., Lukka M.,  
 RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
 RA Manninen V., Ehnholm C.;  
 RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
 RT asparagine to serine substitution at residue 127.";  
 RL Biochim. Biophys. Acta 1138:27-33(1992).  
 RN [14]  
 RP VARIANT A-IV\*5.  
 RX MEDLINE=93138374; PubMed=1487136;  
 RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
 RA Ferrell R.E., Pollitzer W.S.;  
 RT "Molecular basis of a unique African variant (A-IV 5) of human  
 RT apolipoprotein A-IV and its significance in lipid metabolism.";  
 RL Genet. Epidemiol. 9:379-388(1992).  
 RN [15]  
 RP VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
 RX MEDLINE=95245341; PubMed=7728150;  
 RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
 RA Csaszar A.;  
 RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
 RT frequencies, effect on lipid levels, and sequence of two new  
 RT variants.";  
 RL Hum. Mutat. 5:58-65(1995).  
 RN [16]  
 RP VARIANTS FCHL SEATTLE SER-161; LEU-178 AND GLN-264.  
 RX MEDLINE=97114287; PubMed=8956036;  
 RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;  
 RT "Two novel apolipoprotein A-IV variants in individuals with familial  
 RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
 RT activity.";  
 RL Hum. Mutat. 8:319-325(1996).  
 RN CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND  
 CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN  
 CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR  
 CC COMPONENT OF HDL AND CHYLOMICRONS.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.  
 CC SECRETED IN PLASMA.

```
CC -!- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -!- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV*0
CC TO APOA-IV*7). APOA-IV*1 IS THE MAJOR ALLELE (90%), IV*2 IS ALSO
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.
CC -!- DISEASE DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).
CC -!- SIMILARITY: BELONGS TO THE APOAI / APOA4 / APOE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outpost -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M14642; AAA51745.1; -
DR EMBL; X13629; CAA31955.1; -
DR EMBL; M14566; AAA51748.1; -
DR EMBL; J02758; AAA96731.1; -
DR EMBL; M13654; AAA51744.1; -
DR PIR; A26481; LPNUA4.
DR PIR; A24449; A24449.
DR PIR; A29330; A29330.
DR PIR; A26280; A26280.
DR PIR; S02715; S02715.
DR HSSP; P02649; INFO.
DR SWISS-2DPAGE; P06727; HUMAN.
DR MIM; 107690; -
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;
KW Disease mutation; Polymorphism.
FT SIGNAL 1 20
FT CHAIN 21 396 APOLIPOPROTEIN A-IV.
FT DOMAIN 33 330 13 x 22 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 33 54 1.
FT REPEAT 60 81 2.
FT REPEAT 82 103 3.
FT REPEAT 115 136 4.
FT REPEAT 137 158 5.
FT REPEAT 159 180 6.
FT REPEAT 181 202 7.
FT REPEAT 203 224 8.
FT REPEAT 225 246 9.
FT REPEAT 247 268 10.
FT REPEAT 269 286 11.
FT REPEAT 287 308 12.
FT REPEAT 309 330 13.
FT DOMAIN 372 389 GLU/GLN-RICH.
FT VARIANT 13 V -> M (IN APOA-IV*1D).
FT FT 44 /FTId=VAR_000626.
FT FT 44 E -> K (IN BUDAPEST-2).
FT VARIANT 147 147 /FTId=VAR_000627.
FT FT 147 N -> S (IN APOA-IV*1B).
FT VARIANT 161 161 /FTId=VAR_000628.
FT FT 161 A -> S (IN SEATTLE-3; IN FCHL).
FT FT 161 /FTId=VAR_000629.

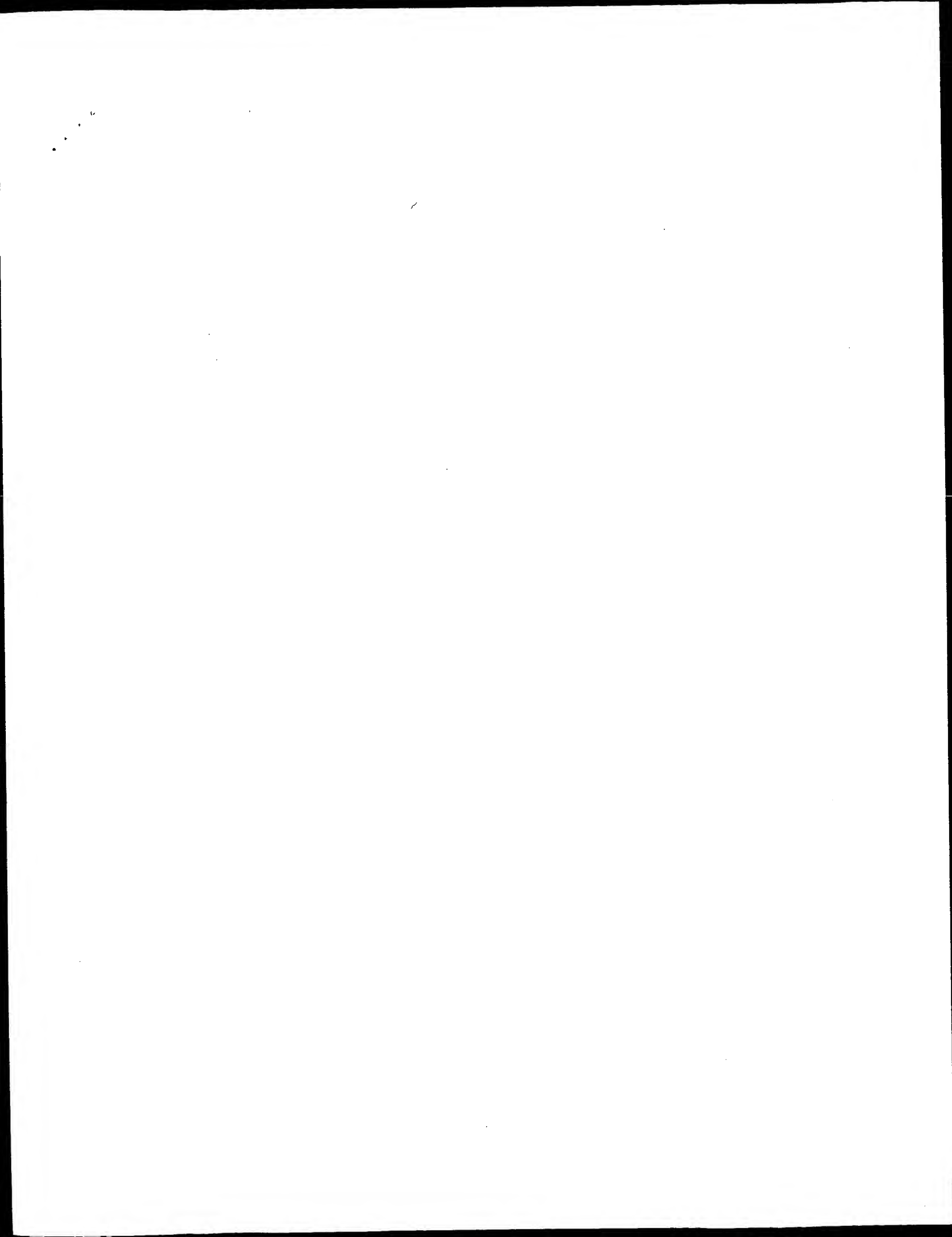
Query Match 19.7%; Score 92; DB 1; Length 396;
Best local Similarity 27.2%; Pred. No. 0.065;
Matches 22; Conservative 19; Mismatches 40; Indels 0; Gaps 0;
```



Sun Sep.22 12:10:20 2002

us-09-803-918a-2\_copy\_25\_113.rsp





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:21:38 ; Search time 111.17 Seconds  
(without alignments)  
138.496 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_113  
Perfect score: 468  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....EFWDNLEKETEGLRQEMSKD 89

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370	79.1	264	11 Q9Z2L4	Q9Z2L4 mesocricetu
2	344	73.5	258	11 Q09054	Q09054 rattus norv
3	344	73.5	258	11 Q08877	Q08877 rattus norv
4	343	73.3	263	11 Q09042	Q09042 mus musculu
5	343	73.3	263	11 Q08855	Q08855 mus musculu
6	292	62.4	241	6 Q9TS49	Q9TS49 erinaceus e
7	217	46.4	82	6 Q29248	Q29248 sus scrofa
8	194	41.5	56	6 Q02762	Q02762 ovis aries
9	141	30.1	26	4 Q9UCT8	Q9UCT8 homo sapien
10	106	22.6	34	13 Q9PRR6	Q9PRR6 anser anser
11	101	21.6	30	11 Q9QV04	Q9QV04 rattus sp.
12	99	21.2	174	13 Q9DF03	Q9DF03 gillithichys
13	84.5	18.1	366	13 Q09601	Q09601 gallus gall
14	81	17.3	435	11 Q01488	Q01488 mus musculu
15	78.5	16.8	687	5 Q95095	Q95095 caenorhabdi
16	78.5	16.8	2697	5 Q01438	Q01438 caenorhabdi

17	77	16.5	395	11 Q9DBN0	Q9dbn0 mus musculu
18	77	16.5	395	11 Q91XF8	Q91xf8 mus musculu
19	74	15.8	313	6 Q9GLC0	Q9glc0 tupia glis
20	73	15.6	491	5 Q45988	Q45988 caenorhabdi
21	72	15.4	102	6 Q29258	Q29258 sus scrofa
22	71.5	15.3	625	16 Q98Q95	Q98q95 mycoplasma
23	71.5	15.3	1687	10 Q9XGN9	Q9xgn9 oryza sativ
24	71	15.2	562	12 Q67018	Q67018 influenza a
25	70.5	15.1	2752	5 Q9BJY0	Q9bjy0 plasmodium
26	70	15.0	328	17 Q9HSV1	Q9hsv1 halobacteri
27	69.5	14.9	452	2 Q9F1X5	Q9f1x5 lactobacilli
28	69.5	14.9	545	4 Q96C12	Q96c12 homo sapien
29	69	14.7	272	4 Q75663	Q75663 homo sapien
30	69	14.7	386	10 Q9LHD0	Q9lhd0 arabidopsis
31	69	14.7	926	17 Q9HM79	Q9hm79 halobacteri
32	69	14.7	1560	9 Q64282	Q64282 streptococc
33	68.5	14.6	160	4 Q9UP14	Q9up14 homo sapien
34	68.5	14.6	329	11 Q9D495	Q9d495 mus musculu
35	68.5	14.6	391	4 Q9Y6S7	Q9y6s7 homo sapien
36	68.5	14.6	633	16 Q9CN49	Q9cn49 pasteurilla
37	68.5	14.6	1205	5 Q9VM93	Q9vm93 drosophila
38	68	14.5	179	11 Q9DAK6	Q9dak6 mus musculu
39	68	14.5	179	11 Q9WUD8	Q9wud8 mus musculu
40	68	14.5	244	4 Q13784	Q13784 homo sapien
41	68	14.5	509	5 Q95SK3	Q95sk3 drosophila
42	68	14.5	4151	5 Q96936	Q96936 drosophila
43	68	14.5	5201	5 Q9U479	Q9u479 drosophila
44	68	14.5	5385	5 Q9V6V3	Q9v6v3 drosophila
45	68	14.5	5496	5 Q9V6V2	Q9v6v2 drosophila

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	264 AA.
ID	Q9Z2L4			
AC	Q9Z2L4			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	APOLIPROTEIN A-I.			
GN	APOL.			
OS	Mesocricetus auratus (Golden hamster).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Mesocricetus.			
OX	NCBI_TaxID=10036;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;			
RX	MEDLINE=99061559; PubMed=9843713;			
RA	Wu J.Y.J., Reaves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;			
RT	"Zinc deficiency decreases plasma level and hepatic mRNA abundance of			
RT	apolipoprotein A-I in rats and hamsters."			
RL	Am. J. Physiol. 275:C1516-C1525(1998).			
DR	EMBL; AF046919; AAC98484.1; -.			
DR	HSSP; P02647; IAV1.			
DR	InterPro; IPR000074; Apolipoprotein.			
DR	Pfam; PF01442; Apolipoprotein; 1.			
KW	Lipoprotein.			
SQ	SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;			

Query Match 79.1%; Score 370; DB 11; Length 264;  
Best Local Similarity 77.3%; Pred. No. 4.1e-31;  
Matches 68; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY	2	EPPOSFMDRVKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLLDNWDSTVTSFKLR 61
DB	25	DDPQTPWDRVKDFATVYVDVAVKDSGREYVSQFETSAIGKQLNLLENWDITGSTVGRLO 84
QY	62	EQLGPVTQEFWDNLEKETEGLRQEMSKD 89









```

QY 69 QEFWDNLEKETEGLRQEMSKD 89
      : | | | | : : | : |
Db 85 VQLSGHLAKETERVKEEIKKE 105

RESULT 15
Q95Q95 PRELIMINARY; PRT; 687 AA.
AC Q95Q95;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 78.2 KDA PROTEIN.
GN B0261.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Scheet P., Gattung S.;
RT "The sequence of C. elegans cosmid B0261.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U97016; AAK84625.1; -.
KW Hypothetical protein.
SQ SEQUENCE 687 AA; 78207 MW; D8EE27A683755777 CRC64;

Query Match 16.8%; Score 78.5; DB 5; Length 687;
Best Local Similarity 44.7%; Pred. No. 8.1;
Matches 21; Conservative 4; Mismatches 17; Indels 5; Gaps 2.

QY 18 YVDVLKDSGRDYSQFEGSALGKQLNLKLDNMWDSVTFSKLRQL 64
      | | | | | | | | | | | | | | | | | | | | | |
Db 240 YSSDLKEAGR-YVQAFESSGNVKDLN----QAWETIYCSVFKKLRDQL 281

Search completed: September 22, 2002, 12:21:39
Job time: 1119 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:05:16 ; Search time 139.03 Seconds  
(without alignments)  
71.104 Million cell updates/sec

Title: us-09-803-918a-2\_COPY\_25\_113  
Perfect score: 468  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....EFWDNLEKETEGLRQMSKD 89

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A\_Geneseq\_032802:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	243	9 AAP81082	Sequence of mature Apo-lipoprotein AI
2	468	100.0	264	15 AAR56863	Apo-lipoprotein AI
3	468	100.0	264	15 AAR56864	Assumed human apol
4	468	100.0	267	7 AAP61079	Entire human prepr
5	468	100.0	267	9 AAP82128	Sequence of apo AI
6	468	100.0	267	14 AAR34032	Human apo A-I incl
7	468	100.0	267	16 AAR72705	Human apolipoprote
8	468	100.0	267	18 AAW08602	Human apolipoprote
9	468	100.0	267	20 AAY18675	Full length Apo-AI
10	468	100.0	267	22 AAB47620	Recombinant human
11	468	100.0	268	9 AAP80668	

12	468	100.0	299	22 AAU33170	Novel human secret
13	415	88.7	154	22 AAO12095	Human polypeptide
14	338	72.2	151	22 AAO02278	Human polypeptide
15	322.5	68.9	120	22 AAU30267	Novel human secret
16	322.5	68.9	120	22 AAU30469	Novel human secret
17	322.5	68.9	166	22 AAU28372	Novel human secret
18	322.5	68.9	244	22 AAU28184	Novel human secret
19	295	63.0	221	22 AAU29835	Novel human secret
20	247	52.8	119	22 AAU30468	Novel human secret
21	246.5	52.7	318	22 AAU30268	Novel human secret
22	183.5	39.2	359	22 AAU30470	Novel human secret
23	116	24.8	21	10 AAP90956	Apo AI epitope. S
24	95.5	20.4	377	14 AAR39502	Human apoAIV mutei
25	92	19.7	183	14 AAR39482	Human apoAIV mutei
26	92	19.7	333	14 AAR39481	Human apoAIV mutei
27	92	19.7	333	14 AAR39488	Human apoAIV mutei
28	92	19.7	333	14 AAR39490	Human apoAIV mutei
29	92	19.7	333	14 AAR39495	Human apoAIV mutei
30	92	19.7	333	14 AAR39497	Human apoAIV mutei
31	92	19.7	337	14 AAR39492	Human apoAIV mutei
32	92	19.7	337	14 AAR39494	Human apoAIV mutei
33	92	19.7	342	14 AAR39487	Human apoAIV mutei
34	92	19.7	342	14 AAR39489	Human apoAIV mutei
35	92	19.7	342	14 AAR39491	Human apoAIV mutei
36	92	19.7	342	14 AAR39496	Human apoAIV mutei
37	92	19.7	342	14 AAR39498	Human apoAIV mutei
38	92	19.7	346	14 AAR39493	Human apoAIV mutei
39	92	19.7	377	14 AAR39443	Human apolipoprote
40	92	19.7	377	14 AAR39480	Human apoAIV mutei
41	92	19.7	377	14 AAR45242	Human apoAIV mutei
42	92	19.7	377	14 AAR45243	Human apoAIV mutei
43	92	19.7	377	14 AAR45244	Human apoAIV mutei
44	92	19.7	396	22 AAB90664	Human secreted pro
45	91	19.4	16	20 AAY27065	Peptide Seq ID No:

ALIGNMENTS

RESULT 1	
AAP81082	Sequence of mature human apolipoprotein AI (apoAI).
ID AAP81082 standard; protein; 243 AA.	
AC AAP81082;	
XX	
XX	14-JAN-1991 (first entry)
XX	
DE	Sequence of mature human apolipoprotein AI (apoAI).
XX	
KW	Atherosclerosis; therapy; cardiovascular disease.
XX	
OS	Homo sapiens.
XX	
PN	W08803166-A.
XX	
PD	05-MAY-1988.
XX	
PF	21-OCT-1987; 87WO-EP00621.
XX	
PR	23-OCT-1986; 86GB-0025435.
XX	
PA	(FARM ) FARMITALIA C ERBA SPA.
XX	
PI	Lorenzetti R, Monaco L, Soria M, Palomba R, Isacchi A, Sarmientos P;
XX	
DR	WPI; 1988-133240/19.
DR	N-PSDB; AAN80243.
XX	
PT	Recombinant human apo-lipoprotein AI -
PT	used to lower plasma cholesterol and/or tri glyceride levels and
PT	to combat atherosclerosis and cardiovascular diseases
XX	
PS	Disclosure; Fig 1; 51pp; English.





```
XX 07-OCT-1991 (first entry)
DT
XX
DE Assumed human apolipoprotein A-1 derivative gene product.
XX
KM Hyperlipaemia; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN JP61096998-A.
XX
PD 15-MAY-1986.
XX
PF 16-OCT-1984; 84JP-0216988.
XX
PR 16-OCT-1984; 84JP-0216988.
XX
PA (MITU ) MITSUBISHI CHEM IND KK.
XX
DR WPI; 1986-165025/26.
DR N-PSDB; AAN60886.
XX
PT Human apo:lipoprotein A-1 (deriv.) prepn. - by providing DNA
PT fragment in cloning site downstream of expression vector promoter
PT and introducing into host microorganism.
XX
PS Disclosure; Fig 2; 9pp; Japanese.
XX
CC The human apolipoprotein may be produced by a suitable transformed
CC host, it is effective in treating hyperlipaemia and arteriosclerosis.
XX
SQ Sequence 267 AA;

Query Match 100.0%; Score 468; DB 7; Length 267;
Best Local Similarity 100.0%; Pred. NO. 7.3e-43;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSL 60
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfsl 84
QY 61 REQLGPVTQEFWMDNLEKETEGLRQEMSKD 89
Db 85 reqlgpvtqefwmdnleketeglrqemskd 113

RESULT 5
AAP82128
ID AAP82128 standard; protein; 267 AA.
XX
AC AAP82128;
XX
DT 24-OCT-1990 (first entry)
XX
DE Entire human preproapoprotein A1.
XX
KM human preproapoprotein A1; high density lipoprotein deficiency; ss.
XX
OS synthetic.
XX
FH Key
FH Peptide 1..18 Location/Qualifiers
FT /label=precursor
FT Peptide 19..24
FT /label=propeptide
FT Protein 25..267
FT /label=mature apoprotein
PN EP293357-A.
XX
PD 30-NOV-1988.
XX
```

```
PF 24-MAY-1988; 88EP-0870095.
XX
XX 28-MAY-1987; 87GB-0012540.
PR
XX
PA (UNIO ) UCB SA.
XX
PI Bollen A, Gobert J, Wulfert E;
XX
DR WPI; 1988-339891/48.
DR N-PSDB; AAN82064.
XX
XX New DNA encoding human preproapoprotein A1 -
PT modified to eliminate hairpin structures
PT
XX Disclosure; ; P; French.
XX
CC The cDNA 878bp fragment encoding preproapoprotein A1 was detected
CC in clone PUB1609 derived from human liver cells.
CC See also AAN81258.
XX
SQ Sequence 267 AA;

Query Match 100.0%; Score 468; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. NO. 7.3e-43;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPMDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSL 60
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalgkqlnlkldnwdsvtstfsl 84
QY 61 REQLGPVTQEFWMDNLEKETEGLRQEMSKD 89
Db 85 reqlgpvtqefwmdnleketeglrqemskd 113

RESULT 6
AAR34032
ID AAR34032 standard; Protein; 267 AA.
XX
AC AAR34032;
XX
DT 13-AUG-1993 (first entry)
XX
DE Sequence of apo A1.
XX
KM Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.
XX
OS Homo sapiens.
XX
PN WO9307165-A.
XX
PD 15-APR-1993.
XX
PF 09-OCT-1992; 92WO-US08634.
XX
PR 09-OCT-1991; 91US-0774633.
PR 08-OCT-1992; 92US-0555555.
PR 28-JUN-1992; 92US-0901706.
XX
PA (SCRI ) SCRIPPS RES INST.
XX
PI Curtiss IK, Koduri KR, Smith RS, Witztum JL, Young SG;
XX
DR WPI; 1993-134378/16.
DR N-PSDB; AAO40030.
XX
PT Polypeptide mimic of native apo B-100 and native apo A-I - useful
PT in assays for LDL and HDL in plasma samples
XX
PS Claim 19; Pages 105-106; 137pp; English.
XX
CC The inventors claim a portion of the polypeptide contg. apo B-100
```



XX New variant of human apoA-1 with Cys at position 151 - has  
PT anti-atherogenic activity for treatment and prevention of  
PT cardiovascular disease

PS Claim 2; Page -: 58pp; French.

CC This is the amino acid sequence of a human apolipoprotein A-1 variant  
CC designated the "Paris" variant which has a Cys replacing the Arg residue  
CC at position 151. The substitution is generated by a mutation of  
CC the C nucleotide at position 523 in the wild type gene to a T residue,  
CC changing the encoded residue from an Arg to a Cys. The gene was isolated  
CC from a patient with an unusual pattern of serum lipids i.e. low levels of  
CC apoA-1 and high density lipoprotein (HDL)-cholesterol and high  
CC triglycerides, but showing symptoms of atherosclerosis. The new variant  
CC protein has anti-atherogenic activity so is useful for treatment and  
CC prevention of cardiovascular diseases such as atherosclerosis,  
CC restenosis, myocardial infarct, angina, sudden death, stroke and cardiac  
CC decompensation, or more generally any condition involving genetic or  
CC metabolic deficit of apoA-1.  
CC Note: this sequence is not given in the specification but is generated  
CC from the wild type apoA-1 gene disclosed in the specification and has  
CC the appropriate amino acid changed.

XX Sequence 267 AA;

Query Match 100.0%; Score 468; DB 18; Length 267;  
Best Local Similarity 100.0%; Pred. No. 7.3e-43;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDIATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 25 deppqspwdrvkdiatvyvdvlkdsgrdysqfegsalsgkqlnlkldnwdsvststfskl 84

QY 61 REQLGPVTQEFWMDNLEKETEGLRQEMSKD 89  
Db 85 reqlgpvtqgefwdnleketeglrqemskd 113

#### RESULT 9

AAV18675  
ID AAV18675 standard; Protein; 267 AA.

XX AAV18675;

DT 09-JUL-1999 (first entry)

DE Human apolipoprotein AI protein sequence.

KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.

OS Homo sapiens.

PN WO916409-A2.

PD 08-APR-1999.

PF 28-SEP-1998; 98WO-US20329.

PR 29-SEP-1997; 97US-0940136.

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (DUFO/) DUFOURCQ J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;

PI Sekul R;

XX WPI; 1999-254921/21.  
DR N-PSDB; AAX55971.

XX Nucleic acid encoding apolipoprotein A-I agonist peptides

PS Example; Fig 1; 232pp; English.

CC The present invention describes a nucleic acid (A) encoding an  
CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,  
CC which forms an amphipathic alpha-helix in presence of lipids. (A),  
CC optionally as a complex with lipids, and host cells that contain (A),  
CC are useful for gene therapy, or prevention, of diseases associated with  
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I  
CC deficiency, hypertiglyceridemia and metabolic syndrome, also to treat  
CC endotoxemia (septic shock). Host cells containing (A) can also be used  
CC to study the role of apoA-I in lipid metabolism. (B) can be used  
CC diagnostically, e.g. to measure serum HDL (particularly its  
CC subpopulation involved in retrograde cholesterol transport) and for  
CC imaging the circulatory system or HDL accumulations at fatty streaks.  
CC The present sequence represents human apoA-I.

XX Sequence 267 AA;

Query Match 100.0%; Score 468; DB 20; Length 267;  
Best Local Similarity 100.0%; Pred. No. 7.3e-43;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDIATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
Db 25 deppqspwdrvkdiatvyvdvlkdsgrdysqfegsalsgkqlnlkldnwdsvststfskl 84

QY 61 REQLGPVTQEFWMDNLEKETEGLRQEMSKD 89  
Db 85 reqlgpvtqgefwdnleketeglrqemskd 113

#### RESULT 10

AAB47620  
ID AAB47620 standard; Protein; 267 AA.

XX AAB47620;

DT 21-JAN-2002 (first entry)

DE Full length Apo-A1.

KW Apolipoprotein; Apo-A1; Apo-A-I fragment T-cell activation inhibitor;  
KW APTI; monocyte; IL-1; interleukin 1; TNF; tumour necrosis factor;  
KW acute pancreatitis; Alzheimer's disease; asthma; cancer; fever;  
KW inflammatory bowel disease; ischemia; multiple sclerosis; osteoporosis;  
KW Parkinson's disease; psoriasis; probe.

OS Homo sapiens.

EH Key Location/Qualifiers

FT Binding-site 44..65

FT Binding-site 220..241

FT Domain 74..111

FT Binding-site 149..219

FT Domain 99..120

FT Domain 99..143

FT Domain 66..120

FT Domain 66..120

/label= Helical lipid binding domain  
/label= Helical lipid binding domain  
/label= Helical lipid binding domain  
/label= Receptor binding domain  
/label= Major antigenic epitope domain  
/label= Hinged domain  
/label= Phylogenetically conserved domain



FT Domain 90..111 /note= "Involved in lectin-cholesterol acyltransferase activity"  
FT  
FT Domain 44..65 /label= Amphipathic helix  
FT Domain 66..98 /label= Amphipathic helix  
FT Domain 99..120 /label= Amphipathic helix  
FT Domain 121..142 /label= Amphipathic helix  
FT Domain 143..164 /label= Amphipathic helix  
FT Domain 165..208 /label= Amphipathic helix  
FT Domain 209..219 /label= Amphipathic helix  
FT Domain 220..241 /label= Amphipathic helix  
FT Peptide 25..194 /label= APTI  
FT Peptide /note= "18 kD N-terminal fragment"  
FT Peptide 25..144 /label= APTI  
FT Peptide /note= "13 kD N-terminal fragment"  
FT Peptide 156..267 /label= APTI  
FT Peptide /note= "13 kD C-terminal fragment"

WO200168852-A2.

20-SEP-2001.

13-MAR-2001; 2001WO-US07826.

13-MAR-2000; 2000US-189008P.

(AMGE-) AMGEN INC.

Edwards CK, Burger D, Dayer J, Kohno T;

WPI; 2001-596308/67.  
N-PSDB; AAH43623.

Apo-A-I fragment T-cell activation inhibitor (AFTI) polynucleotides, useful for treating, diagnosing, ameliorating diseases associated with IL-1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease and asthma -

Claim 1; Fig 1A; 132pp; English.

This sequence shows full length apolipoprotein (Apo-AI). Fragments of Apo-AI may be used as Apo-A-I fragment T-cell activation inhibitors (AFTI). These fragments are selected from an 18 kD N-terminal fragment (amino acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144) and a 13 kD C-terminal fragment (amino acids 156-267). The AFTI polypeptides and polynucleotides are useful for regulating T-cell mediated activation of monocytes and for treating, diagnosing, ameliorating diseases associated with IL-1 and/or TNF activity. The diseases are acute pancreatitis, Alzheimer's disease, asthma, cancer, fever, inflammatory bowel disease, ischemia, multiple sclerosis, osteoporosis, Parkinson's disease or psoriasis. Numerous examples of other diseases are given in the specification. The AFTI nucleic acids are useful as hybridization probes in diagnostic assays to test for the presence of an AFTI or Apo-A-I DNA in mammalian tissue or bodily fluid samples.

Sequence 267 AA;

Query Match 100.0%; Score 468; DB 22; Length 267;  
Best Local Similarity 100.0%; Pred. No. 7.3e-43;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTFSKL 60  
Db 25 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalsgkqlnlkldnwdsvstfstskl 84  
QY 61 REQLGPVTQEFWDNLEKTEGLRQEMSKD 89  
Db 85 reqlgpvtqefwdnleketeglrqemskd 113  
RESULT 11  
AAP80668  
ID AAP80668 standard; protein; 268 AA.  
XX  
AC AAP80668;  
DT 24-OCT-1990 (first entry)  
DE Recombinant human preproapoprotein A1.  
XX human preproapoprotein A1; high density lipoprotein deficiency; ss.  
XX synthetic.  
OS  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18 /label=precursor  
FT Peptide 19..25 /label=proprotein  
FT Protein 26..268 /label=mature apoprotein A1

EP293357-A.

30-NOV-1988.

24-MAY-1988; 88EP-0870095.

28-MAY-1987; 87GB-0012540.

(UNIO ) UCB SA.

Bollen A, Gobert J, Wulfert E;

WPI; 1988-339891/48.  
N-PSDB; AAN81258.

New DNA encoding human preproapoprotein A1 - modified to eliminate hairpin structures

Claim 1; Page 12; 25pp; French.

Met at posn 19 is inserted as an extra amino acid c.f. wild-type protein. The DNA fragment used to replace the wild-type sequence encoding amino acids -6 to +14 (= Arg 20 to Leu 39) modifies codons corresponding to wild-type amino acids -6, -1, 1, 3, 4, 5, 6, 7, 10, 11 and 14. The changed codons still encode the same amino acids as in the wild-type protein but reduce formation of secondary structures in mRNA. See also AAN82064.

Sequence 268 AA;

Query Match 100.0%; Score 468; DB 9; Length 268;  
Best Local Similarity 100.0%; Pred. No. 7.3e-43;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTFSKL 60  
Db 26 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalsgkqlnlkldnwdsvstfstskl 85

QY 61 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 89  
|  
Db 86 reqlgpvtqefwlnleketeglrqemskd 114

## RESULT 12

AAU33170

ID AAU33170 standard; Protein; 299 AA.

AC AAU33170;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3661.

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -

PS Claim 20; Page 718; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

SQ Sequence 299 AA;

Query Match 100.0%; Score 468; DB 22; Length 299;  
Best Local Similarity 100.0%; Pred. No. 8.4e-43;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSL 60  
|  
Db 57 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalsgkqlnlkldnwdsvststfsl 116

QY 61 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 89  
|  
Db 117 reqlgpvtqefwlnleketeglrqemskd 145

## RESULT 13

AAO12095

ID AAO12095 standard; Protein; 154 AA.

AC AAO12095;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 25987.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

PS Claim 20; SEQ ID NO 25987; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 154 AA;

Query Match 88.7%; Score 415; DB 22; Length 154;  
Best Local Similarity 88.8%; Pred. No. 2.1e-37;  
Matches 79; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 DEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSL 60  
|  
Db 33 deppqspwdrvkdlatvyvdvlkdsgrdyvsqfegsalsgkqlnlkldnwdsvststfsl 92

QY 61 REQLGPVTQEFWDLNLEKETEGLRQEMSKD 89  
|  
Db 93 reqlgpvtqefwlnleketeglrqemskd 121

RESULT 14

AA002278  
ID AA002278 standard; Protein; 151 AA.  
XX  
AC AA002278;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 16170.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
DR N-PSDB; AA182209.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 20; SEQ ID NO 16170; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 151 AA;

Query Match 72.2%; Score 338; DB 22; Length 151;  
Best Local Similarity 95.6%; Pred. No. 5e-29;  
Matches 65; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 22 LKDSGRDYVSQFEGSALGKQLNLKLDNWDSTVSTFSKLRQQLGPTVTOEFWDNLEKETEG 81  
Db 1 lkdsgrdyvsqfegsalcglnklkldnwdsetsfstsklreglpgvtgefwnlekdtdeg 60

QY 82 LRQEMSKD 89  
Db 61 lrqemskd 68

RESULT 15  
AAU030267  
ID AAU030267 standard; Protein; 120 AA.  
XX

AC AAU030267;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #758.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 270; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 120 AA;

Query Match 68.9%; Score 322.5; DB 22; Length 120;  
Best Local Similarity 73.0%; Pred. No. 1.8e-27;  
Matches 65; Conservative 1; Mismatches 0; Indels 23; Gaps 1;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDSTVSTFSKL 60  
Db 25 deppqspwdrvkdlatvyvdvlkdsqk-----dsvtstfskl 61

QY 61 RQQLGPTVTOEFWDNLEKETEGLRQEMSKD 89  
Db 62 reqqlgptvtofwdnleketeqlrqemskd 90

Search completed: September 22, 2002, 12:05:16  
Job time: 291 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:02:52 ; Search time 51.49 Seconds  
(without alignments)  
42.219 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_113  
Perfect score: 468  
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....EFWDNLEKETEGLRQEMSKD 89

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	468	100.0	264	1	US-08-448-606-6 Sequence 6, Appli
2	468	100.0	267	1	US-07-959-946-3 Sequence 3, Appli
3	468	100.0	267	1	US-08-333-577-3 Sequence 3, Appli
4	468	100.0	267	4	US-08-952-796-2 Sequence 2, Appli
5	468	100.0	267	5	PCT-US92-08634-3 Sequence 3, Appli
6	241	51.5	200	4	US-08-952-796-15 Sequence 15, Appli
7	66.5	14.2	105	1	US-08-241-853-11 Sequence 11, Appli
8	66.5	14.2	105	2	US-08-850-917-11 Sequence 11, Appli
9	66.5	14.2	180	4	US-09-153-586-24 Sequence 24, Appli
10	66.5	14.2	317	4	US-08-949-155-6 Sequence 6, Appli
11	65	13.9	200	2	US-08-698-805-6 Sequence 6, Appli
12	63.5	13.6	174	2	US-08-768-964-13 Sequence 13, Appli
13	63.5	13.6	174	3	US-09-005-299-13 Sequence 13, Appli
14	63.5	13.6	174	4	US-09-515-431-13 Sequence 13, Appli
15	63.5	13.6	199	2	US-08-768-964-12 Sequence 12, Appli
16	63.5	13.6	199	3	US-09-005-299-12 Sequence 12, Appli
17	63.5	13.6	199	4	US-09-515-431-12 Sequence 12, Appli
18	63.5	13.6	238	2	US-08-768-964-7 Sequence 7, Appli
19	63.5	13.6	238	3	US-09-005-299-7 Sequence 7, Appli
20	63.5	13.6	263	4	US-09-515-431-7 Sequence 7, Appli
21	63.5	13.6	263	2	US-08-768-964-2 Sequence 2, Appli
22	63.5	13.6	263	3	US-09-005-299-2 Sequence 2, Appli
23	63.5	13.6	263	4	US-09-515-431-2 Sequence 2, Appli
24	63	13.5	348	1	US-08-229-781-50 Sequence 50, Appli
25	63	13.5	348	1	US-08-630-918-50 Sequence 50, Appli
26	63	13.5	348	4	US-09-004-422-50 Sequence 50, Appli
27	63	13.5	491	2	US-08-687-916-23 Sequence 23, Appli

28	63	13.5	491	4	US-09-138-614-23	Sequence 23, Appli
29	63	13.5	495	2	US-08-687-916-22	Sequence 22, Appli
30	63	13.5	495	4	US-09-138-614-22	Sequence 22, Appli
31	62.5	13.4	649	3	US-08-996-441B-109	Sequence 109, App
32	62.5	13.4	649	3	US-08-993-722A-109	Sequence 109, App
33	62.5	13.4	649	3	US-08-993-170A-109	Sequence 109, App
34	62.5	13.4	649	3	US-08-993-775B-109	Sequence 109, App
35	62	13.2	796	4	US-08-961-083-56	Sequence 56, Appli
36	61.5	13.1	98	2	US-08-479-078-7	Sequence 7, Appli
37	61.5	13.1	505	4	US-08-426-509A-17	Sequence 17, Appli
38	61.5	13.1	505	5	PCT-US95-05008-17	Sequence 17, Appli
39	60	12.8	620	4	US-09-269-731-6	Sequence 6, Appli
40	60	12.8	729	1	US-08-070-165F-6	Sequence 6, Appli
41	60	12.8	729	2	US-08-885-418-6	Sequence 6, Appli
42	60	12.8	731	1	US-08-070-165F-10	Sequence 10, Appli
43	60	12.8	731	2	US-08-885-418-10	Sequence 10, Appli
44	59.5	12.7	372	3	US-08-918-249-2	Sequence 2, Appli
45	59.5	12.7	372	3	US-08-918-249-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-08-448-606-6  
; Sequence 6, Application US/08448606  
; Patent No. 5721114  
; GENERAL INFORMATION:  
; APPLICANT: Abrahams n, Lars  
; APPLICANT: Holmgren, Erik  
; APPLICANT: Kalder n, Christina  
; APPLICANT: Lake, Mats  
; APPLICANT: Mikaelsson, sa  
; APPLICANT: Sejlitz, Torsten  
; TITLE OF INVENTION: Expression System For Producing  
; TITLE OF INVENTION: Apolipoprotein AI-M  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pollock, Vande Sande & Priddy  
; STREET: 1990 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,606  
; FILING DATE: 25-AUG-1995  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/SE93/01061  
; FILING DATE: 09-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9203753-0  
; FILING DATE: 11-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Amernick, Burton A.  
; REGISTRATION NUMBER: 24,852  
; REFERENCE/DOCKET NUMBER: 0151/00121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)331-7111  
; TELEFAX: (202) 293-6229  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-448-606-6

Query Match	100.0%;	Score 468;	DB 1;	Length 264;
Best Local Similarity	100.0%;	Pred. No. 1.1e-46;		
Matches	89;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	1	DEPPQSPMDRVKDLATVYVVDYVLKDSGRDYVSQFEFGSALGKQLNLKLLDNWDSVTSFE	SKL	60
DB	22	DEPPQSPMDRVKDLATVYVVDYVLKDSGRDYVSQFEFGSALGKQLNLKLLDNWDSVTSFE	SKL	81
QY	61	REOLGPVTOEFWMDNLEKETEGLRQEMSKD		89
DB	82	REOLGPVTOEFWMDNLEKETEGLRQEMSKD		110

```

RESULT      2
US-07-959-946-3
; Sequence 3, Application US/07959946
; Patent No. 5408038
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Wiltzium, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Suter &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5408038th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,946
; FILING DATE: 19921008
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-959-946-3

```

Query Match	100.0%;	Score 468;	DB 1;	Length 267;
Best Local Similarity	100.0%;	Pred. No. 1.1e-46;		
Matches	89;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

```

Db      85 REQLGPTQEEFWDNLEKETEGRLQEMSKD 113

```

```

RESULT      3
; US-08-333-577-3
; Sequence 3, Application US/08333577
; Patent No. 5786206
;
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Wiltzium, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,577
;
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-333-577-3

```

Query Match	100.0%;	Score 468;	DB 1;	Length 267;
Best Local Similarity	100.0%;	Pred. No. 1.1e-46;		
Matches 89; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	DEFPQSPWDRVKDLATFYVYVDYLLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL	60
Db	25	DEFPQSPWDRVKDLATFYVYVDYLLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL	84
QY	61	REOLGPVTOEFWDNLEKTEGLROEMSKD	89
Db	85	REOLGPVTOEFWDNLEKTEGLROEMSKD	113

```

RESULT 4
US-08-952-796-2
; Sequence 2, Application US/08952796
; Patent No. 6258596
; GENERAL INFORMATION:
; APPLICANT: BENOIT, Patrick
; APPLICANT: BRUCKERT, Eric
; APPLICANT: DENEFFE, Patrice
; APPLICANT: DUBERGER, Nicolas
; APPLICANT: FRUCHART, Jean-Charles

```

APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerrard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehlner Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-2

Query Match 100.0%; Score 468; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-46;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
QY 61 REQLGPVTQEFWMDLKEKETEGLRQEMSKD 89  
Db 85 REQLGPVTQEFWMDLKEKETEGLRQEMSKD 113

RESULT 5  
PCT-US92-08634-3

; Sequence 3, Application PC/TUS9208634  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 North Stetson, Suite 4700

CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08634  
FILING DATE: 19921009  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-08634-3

Query Match 100.0%; Score 468; DB 5; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.1e-46;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60  
Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84  
QY 61 REQLGPVTQEFWMDLKEKETEGLRQEMSKD 89  
Db 85 REQLGPVTQEFWMDLKEKETEGLRQEMSKD 113

RESULT 6  
US-08-952-796-15

; Sequence 15, Application US/08952796  
; Patent No. 6258596  
; GENERAL INFORMATION:  
; APPLICANT: BENOIT, Patrick  
; APPLICANT: BRUCKERT, Eric  
; APPLICANT: DENEFFLE, Patrice  
; APPLICANT: DUBERGER, Nicolas  
; APPLICANT: FRUCHART, Jean-Charles  
; APPLICANT: LUC, Gerald  
; APPLICANT: TURPIN, Gerrard  
; APPLICANT: ASSMANN, Gerd  
; APPLICANT: FUNKE, Harald  
; TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, Mailstop 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30



;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/952,796  
;; FILING DATE:  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/06061  
;; FILING DATE: 22-MAY-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO FR96/00747  
;; FILING DATE: 20-MAY-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fehner Esq., Paul F.  
;; REGISTRATION NUMBER: 35,135  
;; REFERENCE/DOCKET NUMBER: ST95031-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 200 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-952-796-15

Query Match 51.5%; Score 241; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 1.7e-20;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 LKLDNWDVSTSTFSKRLQGLPVTQEFWDNLEKETEGLRQEMSKD 89  
1 LKLDNWDVSTSTFSKRLQGLPVTQEFWDNLEKETEGLRQEMSKD 46

## RESULT 7

US-08-241-853-11  
; Sequence 11, Application US/08241853  
; Patent No. 5693488

;; GENERAL INFORMATION:

;; APPLICANT: Fang, Kathy S.

;; APPLICANT: Hanafusa, Hidesaburo

;; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE

;; TITLE OF INVENTION: AND METHODS OF USE THEREOF

;; NUMBER OF SEQUENCES: 37

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Klauber & Jackson

;; STREET: 411 Hackensack Avenue

;; CITY: Hackensack

;; STATE: New Jersey

;; COUNTRY: USA

;; ZIP: 07601

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/241,853

;; FILING DATE: 12-MAY-1994

;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Jackson Esq., David A.

;; REGISTRATION NUMBER: 26,742

;; REFERENCE/DOCKET NUMBER: 600-1-078

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 201 487-5800

;; TELEFAX: 201 343-1684

;; TELEX: 133521

;; INFORMATION FOR SEQ ID NO: 11:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 105 amino acids

;; TYPE: amino acid

;;  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; US-08-241-853-11

Query Match 14.2%; Score 66.5; DB 1; Length 105;  
Best Local Similarity 27.7%; Pred. No. 1.3;  
Matches 26; Conservative 11; Mismatches 28; Indels 29; Gaps 4;

QY 8 WDRYKDLATVYVDLKD-----SGRDYVSGFEQS--ALGK 40  
12 WDQKDLAFAFCKKNLQDAENFFQFGDADDLKAWLQDAHRLLSGED-VGQDEGATRALGK 70

QY 41 QLNKLDNWDVSTSTFSKRLQGLPVTQEFWDN 74  
Db 71 K-HKDFLEELLESRGVMEKLEQQAQGFPEEFKRS 103

## RESULT 8

US-08-850-917-11  
; Sequence 11, Application US/08850917  
; Patent No. 5854045

;; GENERAL INFORMATION:

;; APPLICANT: Fang, Kathy S.

;; APPLICANT: Hanafusa, Hidesaburo

;; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE

;; TITLE OF INVENTION: AND METHODS OF USE THEREOF

;; NUMBER OF SEQUENCES: 37

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Klauber & Jackson

;; STREET: 411 Hackensack Avenue

;; CITY: Hackensack

;; STATE: New Jersey

;; COUNTRY: USA

;; ZIP: 07601

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/850,917

;; FILING DATE: 02-MAY-1997

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/241,853

;; FILING DATE: 12-MAY-1994

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Jackson Esq., David A.

;; REGISTRATION NUMBER: 26,742

;; REFERENCE/DOCKET NUMBER: 600-1-078

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 201 487-5800

;; TELEFAX: 201 343-1684

;; TELEX: 133521

;; INFORMATION FOR SEQ ID NO: 11:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 105 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

;; ANTI-SENSE: NO

;; US-08-850-917-11

Query Match 14.2%; Score 66.5; DB 2; Length 105;  
Best Local Similarity 27.7%; Pred. No. 1.3;  
Matches 26; Conservative 11; Mismatches 28; Indels 29; Gaps 4;

```

Query Match          14.2%; Score 66.5; DB 4; Length 180;
Best Local Similarity 26.3%; Pred. No. 2.7;
Matches 20; Conservative 14; Mismatches 29; Indels 13; Gaps 2;

QY 23 KDSGRDYSQFEGSALGKQLNLKLD-----NMDSVTSTFSKLRQLGPVTOE 70
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 RDSPRDHYVQFKGLCYTNGTQRIRDVIRYIYNQEEYLRVDSVDVGEYRALTE-LGRPSAE 61
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 71 FWDNLEKETEGLRQEM 86
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 YWNSOKOYLEQTRAEI 77

RESULT 10
US-08-949-155-6
; Sequence 6, Application US/08949155
; Patent No. 6271436
; GENERAL INFORMATION:
; APPLICANT: Piedrahita, Jorge A
; APPLICANT: Bazer, Fuller W
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800

```

```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/027,338
: FILING DATE: 11-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/046,094
: FILING DATE: 09-MAY-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Hibler, David W.
: REGISTRATION NUMBER: 41,071
: REFERENCE/DOCKET NUMBER: TAMR:177
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 317 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-949-155-6

```

```
Query Match      14.28; Score 66.5; DB 4; Length 317;  
Best Local Similarity 21.78; Pred. No. 5.6;  
Matches 20; Conservative 19; Mismatches 50; Indels 3; Gaps 1;  
  
QY   1 DEPP---QSPWDRVKDLATVYVDVLKDSGRDYVSQFECSALGKQLNLKLNDWDSVTSTF 57  
    :| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db   33 EEPKWQSGQPWEQALGRFWMDYLIRWVQSLSDQVQEELLSTKVTOELTEELIEESMEKKAYR 92  
  
QY   58 SKLRQGLGPVTQEFWDNLEKETEGELRQEMSKD 89  
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db   93 EELEAQGLGPVTQETQARLSKELQAQAQARVGAD 124
```

RESULT 11  
US-08-698-805-6  
Sequence 6, Application US/08698805  
Patent No. 5869288  
GENERAL INFORMATION:  
APPLICANT: Chapman, Martin  
APPLICANT: Arruda, L. Karla  
TITLE OF INVENTION: Molecular Cloning of Cockroach  
TITLE OF INVENTION: Allergens, Amino Acid and Nucleotide Sequences Therefore,  
TITLE OF INVENTION: and Recombinant Expression Thereof  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/698,805  
FILING DATE: 16-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/002,510  
FILING DATE: 18-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kelber, Steven B.  
REGISTRATION NUMBER: 30,073  
REFERENCE/DOCKET NUMBER: 494-203-27  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220



APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: No. 6284881e1 Feline Fc Epsilon Receptor Alpha  
TITLE OF INVENTION: Chain Nucleic Acid Molecules, Proteins  
TITLE OF INVENTION: and Uses Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/515,431  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/768,964  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: DI-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/493-7333  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-515-431-13

Query Match 13.6%; Score 63.5; DB 4; Length 174;  
Best Local Similarity 36.6%; Pred. No. 5.7;  
Matches 15; Conservative 10; Mismatches 11; Indels 5; Gaps 2;

QY 14 LATVYVDVLK---DSGRDYVSQFEGSALGKQLNLKLLDNW 50  
DB 49 VTTLTLDIVKAQIRDSG-EYTCQNKGSMLSKPVSLKVFREW 88

RESULT 15  
US-08-768-964-12  
Sequence 12, Application US/08768964  
Patent No. 5958880  
GENERAL INFORMATION:  
APPLICANT: Frank, Glenn R.  
APPLICANT: Porter James P.  
APPLICANT: Rushlow, Keith E.  
APPLICANT: Wassom, Donald L.  
APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: No. 5958880e1 Feline Fc Epsilon Receptor Alpha  
TITLE OF INVENTION: Chain Nucleic Acid Molecules, Proteins  
TITLE OF INVENTION: and Uses Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/768,964  
FILING DATE: December 19, 1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: DI-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/493-7333  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 199 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-768-964-12

Query Match 13.6%; Score 63.5; DB 2; Length 199;  
Best Local Similarity 36.6%; Pred. No. 6.8;  
Matches 15; Conservative 10; Mismatches 11; Indels 5; Gaps 2;

QY 14 LATVYVDVLK---DSGRDYVSQFEGSALGKQLNLKLLDNW 50  
DB 74 VTTLTLDIVKAQIRDSG-EYTCQNKGSMLSKPVSLKVFREW 113

Search completed: September 22, 2002, 12:02:53  
Job time: 253 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:19:45 ; Search time 17.64 Seconds

(without alignments)  
1454.413 Million cell updates/sec

Title: US-09-803-918A-2

Perfect score: 267

Sequence: 1 MKAVALTLAVLFLTGSQARH.....SPKVSFLSAL E EYTKLNTQ 267

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	100.0	267	1	LPHUA1
2	69	25.8	267	1	A26529
3	69	25.8	267	2	JS0079
4	25	9.4	241	2	A24998
5	22	8.2	265	1	LPRB1B
6	22	8.2	266	1	LPRB1Z
7	22	8.2	266	1	LPDGA1
8	21	7.9	264	2	S31394
9	21	7.9	265	2	A46018
10	21	7.9	265	2	A56858
11	21	7.9	265	2	JT0672
12	20	7.5	231	2	JQ0704
13	19	7.1	164	2	S21830
14	12	4.5	259	2	A24700
15	12	4.5	262	2	JC1237
16	12	4.5	264	2	S22420
17	10	3.7	132	2	G83132
18	10	3.7	246	2	A61448
19	10	3.7	264	1	LPCHA1
20	9	3.4	264	2	JC5456
21	8	3.0	20	2	A05313
22	8	3.0	300	2	G83100
23	8	3.0	317	2	S33450
24	8	3.0	387	2	F82692
25	8	3.0	391	2	B40892
26	8	3.0	394	2	A25281
27	8	3.0	395	2	A40892
28	8	3.0	399	2	C40892
29	8	3.0	1950	2	S12332

30	7	2.6	67	2	T05224	hypothetical prote
31	7	2.6	101	2	H87207	probable membrane
32	7	2.6	116	2	E71566	probable S16 ribos
33	7	2.6	132	2	C86709	hypothetical prote
34	7	2.6	139	2	G72710	hypothetical prote
35	7	2.6	150	2	T33928	hypothetical prote
36	7	2.6	150	2	E86711	hypothetical prote
37	7	2.6	155	2	AB0918	conserved hypothet
38	7	2.6	161	2	F91222	hypothetical prote
39	7	2.6	161	2	S30747	hypothetical 17.1k
40	7	2.6	161	2	D86069	hypothetical prote
41	7	2.6	199	2	H95871	conserved hypothet
42	7	2.6	205	2	H84357	hypothetical prote
43	7	2.6	216	2	T37221	hypothetical prote
44	7	2.6	218	2	E85021	hypothetical prote
45	7	2.6	219	2	G65089	hypothetical prote
46	7	2.6	219	2	C91117	probable transcrip
47	7	2.6	219	2	C85962	probable 2-compone
48	7	2.6	219	2	A10888	probable 2-compone
49	7	2.6	219	2	F65018	hypothetical prote
50	7	2.6	219	2	F91042	hypothetical prote
51	7	2.6	219	2	A85887	ethanolamine utili
52	7	2.6	219	2	AE0813	Surf1 family prote
53	7	2.6	228	2	F87670	hypothetical prote
54	7	2.6	231	2	F95267	hypothetical prote
55	7	2.6	244	2	AC3341	precorrin-2 C20-me
56	7	2.6	250	1	A44039	chitinase (EC 3.2.
57	7	2.6	265	2	T46013	hypothetical prote
58	7	2.6	266	2	H83008	N-formylglutamate
59	7	2.6	267	2	T37047	probable nitrate r
60	7	2.6	269	2	C90450	hypothetical prote
61	7	2.6	273	2	S21731	signal recognition
62	7	2.6	291	2	AG0907	probable kinase [i
63	7	2.6	296	2	A95408	hypothetical prote
64	7	2.6	301	2	B98211	amidophosphoribosy
65	7	2.6	301	2	AG3075	amidophosphoribosy
66	7	2.6	314	2	E32307	DNA-directed RNA p
67	7	2.6	316	2	D71375	probable ABC trans
68	7	2.6	322	2	G97396	cysteine synthase
69	7	2.6	322	2	AH2614	cysteine synthase
70	7	2.6	330	2	D64526	hypothetical prote
71	7	2.6	330	2	JC5935	exostose-related p
72	7	2.6	335	2	T29462	hypothetical prote
73	7	2.6	339	2	E82211	conserved hypothet
74	7	2.6	339	2	E97345	oligopeptide ABC t
75	7	2.6	350	2	G01950	hypothetical prote

## ALIGNMENTS

RESULT 1  
LPHUA1  
apolipoprotein A-I precursor [validated] - human  
N:Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor  
C:Species: Homo sapiens (man)  
C:Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C:Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010;  
6197  
R:Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.  
DNA 3, 309-317, 1984  
A:Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for hu  
A:Reference number: A90947; MUID:85026665  
A:Accession: A90947  
A:Molecule type: DNA  
A:Residues: 1-267 <SE1>  
A:Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
A:Accession: B90947  
A:Molecule type: mRNA  
A:Residues: 1-267 <SE2>  
A:Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635  
R:Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis,  
Eur. J. Biochem. 173, 465-471, 1988

A;Title: Sequence and expression of Tangier apoA-I gene.  
 A;Reference number: S02373; MUID:88196137  
 A;Accession: S02373  
 A;Molecule type: DNA  
 A;Residues: 1-267 <MAK>  
 A;Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729  
 R;Shoulders, C.C.; Kornblitt, A.R.; Munro, B.S.; Baralle, F.E.  
 Nucleic Acids Res. 11, 2827-2837, 1983  
 A;Title: Gene structure of human apolipoprotein A1.  
 A;Reference number: A93465; MUID:83220822  
 A;Accession: A93465  
 A;Molecule type: DNA  
 A;Residues: 1-267 <SHO>  
 A;Cross-references: GB:J00098; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:K01519  
 R;Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983  
 A;Title: Isolation and characterization of the human apolipoprotein A-I gene.  
 A;Reference number: A21147; MUID:84016011  
 A;Accession: A21147  
 A;Molecule type: DNA  
 A;Residues: 1-267 <KAR>  
 A;Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768  
 R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.E.  
 Nucleic Acids Res. 12, 3917-3932, 1984  
 A;Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundance.  
 A;Reference number: A93519; MUID:84221405  
 A;Accession: A93519  
 A;Molecule type: mRNA  
 A;Residues: 1-267 <SHA>  
 A;Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753  
 A;Accession: B93519  
 A;Molecule type: DNA  
 A;Residues: 1-24 <SH2>  
 R;Cheung, P.; Chan, L.  
 Nucleic Acids Res. 11, 3703-3715, 1983  
 A;Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.  
 A;Reference number: A93472; MUID:83220772  
 A;Accession: A93472  
 A;Molecule type: mRNA  
 A;Residues: 1-267 <CHE>  
 A;Cross-references: GB:J00098; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:K01519  
 R;Law, S.W.; Brewer Jr., H.B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984  
 A;Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA  
 A;Reference number: A94010; MUID:84119464  
 A;Accession: A94010  
 A;Molecule type: mRNA  
 A;Residues: 1-267 <LAM>  
 A;Cross-references: GB:J00098; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:K01519  
 R;Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983  
 A;Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted  
 A;Reference number: A21118; MUID:83195100  
 A;Accession: A21118  
 A;Molecule type: mRNA  
 A;Residues: 1-24 <ZAN>  
 R;Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A.  
 Biochem. Biophys. Res. Commun. 113, 626-632, 1983  
 A;Title: Human plasma proapoA-I: isolation and amino-terminal sequence.  
 A;Reference number: A90112; MUID:83256553  
 A;Accession: A90112  
 A;Molecule type: protein  
 A;Residues: 19-27 <BRE>  
 R;Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J.  
 Biochem. Biophys. Res. Commun. 80, 623-630, 1978  
 A;Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high cholesterol plasma  
 A;Reference number: A90209; MUID:78123731  
 A;Accession: A90209  
 A;Molecule type: protein  
 A;Residues: 25-57, 'Q', 59-169, 'QQ', 172-267 <BR2>  
 R;Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.  
 J. Clin. Invest. 82, 803-807, 1988  
 A;Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A

A;Reference number: A30516; MUID:88331387  
 A;Accession: A30516  
 A;Molecule type: protein  
 A;Residues: 25-56 <YUI>  
 R;Nichols, W.C.; Dwulet, F.E.; Liepnies, J.; Benson, M.D.  
 Biochem. Biophys. Res. Commun. 156, 762-768, 1988  
 A;Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloidosis  
 A;Reference number: A31582; MUID:89050104  
 A;Accession: A31582  
 A;Molecule type: protein  
 A;Residues: 25-49, 'R', 51-85, 'D', 87-107 <NIC>  
 A;Note: variant sequence from patient with familial amyloidotic polyneuropathy type I  
 R;Manjunath, P.; Marcel, Y.L.; Uma, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A.  
 J. Biol. Chem. 264, 16853-16857, 1989  
 A;Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.  
 A;Reference number: A34409; MUID:89380318  
 A;Accession: A34409  
 A;Molecule type: protein  
 A;Residues: 25-48 <MAN>  
 R;Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988  
 A;Title: Structural requirements of human preproapolipoprotein AI for translocation a  
 A;Reference number: S02737; MUID:89149957  
 A;Accession: S02737  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-6, 'AV', 9, 'LV', 12-29 <STO>  
 A;Note: part of this sequence, including the amino end of the mature protein, was con  
 R;Stoffel, W.; Binczek, E.  
 Biol. Chem. Hoppe-Seyler 372, 481-488, 1991  
 A;Title: Transient expression of wild type and mutant human apolipoprotein AI in COS  
 A;Reference number: S16197; MUID:92029676  
 A;Contents: annotation; extension of studies in reference S02737  
 R;Stoffel, W.; Kruger, E.; Deutzmann, R.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983  
 A;Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processi  
 A;Reference number: A19913; MUID:83236195  
 A;Accession: B19913  
 A;Molecule type: protein  
 A;Residues: 1-6, 'X', 8-13, 'XXX', 17-18, 'XX', 21, 'X', 23-25, 'X', 27-29 <ST2>  
 R;Ehnholm, C.; Bozas, S.E.; Tenkanen, H.; Kirsbaum, L.; Metso, J.; Murphy, B.; Walke  
 Biochim. Biophys. Acta 1086, 255-260, 1991  
 A;Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein  
 A;Reference number: A56815; MUID:92075698  
 A;Accession: A56815  
 A;Molecule type: protein  
 A;Residues: 25-31, 'P', 33 <EHN>  
 A;Experimental source: serum  
 A;Note: sequence extracted from NCBI backbone (NCBIP:69759)  
 A;Note: 32-Trip was also found  
 R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
 Biochemistry 33, 1988-1993, 1994  
 A;Title: Identification of proteins associated with apolipoprotein A-I-containing lip  
 A;Reference number: A54223; MUID:94162201  
 A;Accession: A54223  
 A;Molecule type: protein  
 A;Residues: 25-39 <KUN>  
 R;Mogulievsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.;  
 DNA 8, 429-436, 1989  
 A;Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: p  
 A;Reference number: I39476; MUID:89377481  
 A;Accession: I39476  
 A;Molecule type: mRNA  
 A;Residues: 19-267 <RES>  
 A;Cross-references: GB:M29068; NID:g178774; PIDN:AAA51747.1; PID:g178775  
 R;Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B.  
 J. Biol. Chem. 263, 18530-18536, 1988  
 A;Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by th  
 A;Accession: I39475  
 A;Molecule type: DNA  
 A;Residues: 1-14 <RE2>  
 A;Cross-references: GB:J04066; NID:g178763; PIDN:AAA51746.1; PID:g553183

R;Breslow, J.L.  
Annu. Rev. Biochem. 54, 699-727, 1985  
A;Title: Human apolipoprotein molecular biology and genetic variation.  
A;Reference number: A90042; MUID:85278004  
A;Contents: annotation; review of sequences, variants and gene location  
R;Hoeg, J.M.; Meng, M.S.; Roman, R.; Fairwell, T.; Brewer Jr., H.B.  
J. Biol. Chem. 261, 3911-3914, 1986  
A;Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation  
A;Reference number: A92577; MUID:86140194  
A;Contents: annotation; acylation with palmitate  
A;Note: an undetermined serine or threonine is acylated by fatty acid; the acylating fatty acid is  
R;Law, S.W.; Brewer, H.B.  
J. Biol. Chem. 260, 12810-12814, 1985  
A;Title: Tangier disease: The complete mRNA sequence encoding for preproapo-A-I.  
A;Reference number: I55236; MUID:86008382  
A;Accession: I55236  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-143, 'D', 145-267 <RE3>  
A;Cross-references: GB:M11791; NID:g178776; PIDN:AAA35545.1; PID:g178777  
C;Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestine. oteins (HDL) in plasma.  
C;Genetics:  
A;Gene: GDB:APOA1  
A;Cross-references: GDB:119684; OMIM:107680  
A;Map position: 11q23.3-11q23.3  
A;Introns: 15/1; 67/2  
C;Function:  
A;Description: participates in the reverse transport of cholesterol from tissues to the sterol acyltransferase (LCAT); noncovalently binds and stabilizes prostacyclin (PGI-2)  
C;Superfamily: apolipoprotein A-I  
C;Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipid F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-267/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 100.0%; Score 267; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.2e-253;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFWQODEPQSPWDRVKDLATVYVDVLKDSGRDVSQFEQS 60  
|||||  
Db 1 MKAAYLTAVLFLTGSQARHFWQODEPQSPWDRVKDLATVYVDVLKDSGRDVSQFEQS 60

QY 61 AIGKQLNLKLDNWDVSTSTFSKLRQGLPVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120  
|||||  
Db 61 AIGKQLNLKLDNWDVSTSTFSKLRQGLPVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120

QY 121 VQPYLDDFOKKWQEMELYRQVEPLRAELQEGARQKLHELQEKLSPLGEMDRARAHV 180  
|||||  
Db 121 VQPYLDDFOKKWQEMELYRQVEPLRAELQEGARQKLHELQEKLSPLGEMDRARAHV 180

QY 181 DALRTHLAPYSDELQRRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
|||||  
Db 181 DALRTHLAPYSDELQRRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESEFKVSFLSALAEYTKKLTNQ 267  
|||||  
Db 241 GLLPVLESEFKVSFLSALAEYTKKLTNQ 267

RESULT 2  
A26529  
apolipoprotein A-I precursor - crab-eating macaque  
C;Species: Macaca fascicularis (crab-eating macaque)  
C;Date: 30-Sep-1989 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
C;Accession: A26529; A26627; S23135; A57766  
R;Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marotti, K.R.  
Gene 49, 103-110, 1986  
A;Title: The primary structure of cynomolgus monkey apolipoprotein A-1 deduced from the  
A;Reference number: A26529; MUID:87191989  
A;Accession: A26529

A;Molecule type: mRNA  
A;Residues: 1-267 <POL>  
A;Cross-references: GB:M15411; NID:g342074; PIDN:AAA36834.1; PID:g342075  
R;Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Sartelli, A.L.; Kantor, M.A.; Nicolo  
Biochemistry 26, 1457-1463, 1987  
A;Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (  
A;Reference number: A26627; MUID:87185451  
A;Accession: A26627  
A;Molecule type: protein  
A;Residues: 25-48 <HER>  
R;Murray, R.W.; Marotti, K.R.  
Biochim. Biophys. Acta 1131, 207-210, 1992  
A;Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and cor  
A;Reference number: S23135; MUID:92305062  
A;Accession: S23135  
A;Molecule type: DNA  
A;Residues: 1-12, 'L', 14-267 <MUR>  
A;Cross-references: GB:M83242; NID:g342070; PIDN:AAA36832.1; PID:g342071  
R;Sorci-Thomas, M.; Kearns, M.W.  
J. Biol. Chem. 266, 18045-18050, 1991  
A;Title: Transcriptional regulation of the apolipoprotein A-I gene.  
A;Reference number: A57766; MUID:92011532  
A;Accession: A57766  
A;Molecule type: DNA  
A;Residues: 1-10 <RES>  
A;Cross-references: GB:M69223; NID:g342066; PIDN:AAA36831.1; PID:g553820  
C;Comment: The precursor is synthesized in the liver and small intestine. The propept  
C;Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins  
y promoting cholesterol efflux from tissues and by acting as a cofactor for the lecithin  
C;Genetics:  
A;Introns: 15/1; 67/2  
C;Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PPT>  
F;25-267/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 25.8%; Score 69; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-59;  
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 SKLRQQLGPTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQ 141  
|||||  
Db 82 SKLRQQLGPTQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFOKKWQEMELYRQ 141

QY 142 KVEPLRAEL 150  
|||||  
Db 142 KVEPLRAEL 150

RESULT 3  
JS0079  
apolipoprotein A-I precursor - baboon  
C;Species: Papio sp. (baboon)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 13-Jun-1997  
C;Accession: JS0079  
R;Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; Vandenberg, J.L.  
Gene 74, 483-490, 1988  
A;Title: The baboon gene for apolipoprotein A-I: characterization of a cDNA clone and  
A;Reference number: JS0079; MUID:89232739  
A;Accession: JS0079  
A;Molecule type: mRNA  
A;Residues: 1-267 <HIX>  
A;Experimental source: liver  
A;Comment: This protein is the principal protein component of high density lipoprotei  
C;Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase  
C;Comment: This protein contains a region of repeated amino acids which form amphipat  
C;Genetics:  
A;Gene: apoA1  
A;Superfamily: apolipoprotein A-I  
C;Keywords: HDL; lipid binding; lipoprotein  
F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-267/Product: apolipoprotein A-I #status predicted <LAI>  
F;123-144,145-166,167-188,189-210,211-232,233-254/Region: tandem repeats

Query Match 25.8%; Score 69; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-59;  
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 SKLREQLPVTQEFWDNLEKETEGLRQEMSKDLEVKAKVQPYLDDFQKKQGEEMELYRQ 141  
|||||  
Db 82 SKLREQLPVTQEFWDNLEKETEGLRQEMSKDLEVKAKVQPYLDDFQKKQGEEMELYRQ 141

QY 142 KVEPLRAEL 150  
|||||  
Db 142 KVEPLRAEL 150

RESULT 4  
A24998

apolipoprotein A-I - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 31-Dec-1993

C;Accession: A24998

R;Yang, C.; Yang, T.; Pownall, H.J.; Gotto Jr., A.M.

Eur. J. Biochem. 160, 427-431, 1986

A;Title: The primary structure of apolipoprotein A-I from rabbit high-density lipoprotein

A;Reference number: A24998; MUID:87030294

A;Accession: A24998

A;Molecule type: protein

A;Residues: 1-241 <YAN>

C;Superfamily: apolipoprotein A-I

C;Keywords: HDL; lipid binding; lipoprotein

Query Match 9.4%; Score 25; DB 2; Length 241;  
Best Local Similarity 100.0%; Pred. No. 1.8e-16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EQLGPVTQEFWDNLEKETEGLRQEM 110  
|||||  
Db 60 EQLGPVTQEFWDNLEKETEGLRQEM 84

RESULT 5

LPRB1B

apolipoprotein A-I precursor (clone pRBA-502) - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999

C;Accession: S00230; S20557

R;Pan, T.C.; Hao, Q.L.; Yamin, T.T.; Dai, P.H.; Chen, B.S.; Chen, S.L.; Kroon, P.A.; Cha

Eur. J. Biochem. 170, 99-104, 1987

A;Title: Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit apolipoprotein A-

A;Reference number: S00230; MUID:88082866

A;Accession: S00230

A;Molecule type: mRNA

A;Residues: 1-265 <PAN>

A;Cross-references: EMBL:X06658; NID:g1461; PIDN:CAA29857.1; PID:g1462

A;Note: the authors translated the codon AGC for residue 174 as Arg

A;Accession: S20557

A;Molecule type: DNA

A;Residues: 1-17, 'R', 19-44, 'I', 46-122, 'Y', 124-146, 'V', 148-265 <PAN2>

A;Cross-references: EMBL:X06659; NID:g1459; PIDN:CAA29858.1; PID:g1460

C;Comment: This protein is synthesized in the small intestine.

C;Comment: This protein is a major component of the high density lipoproteins in plasma.

C;Genetics:

A;Introns: 15/1; 66/2

C;Superfamily: apolipoprotein A-I

C;Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-24/Domain: propeptide #status predicted <PRO>

F;25-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 8.2%; Score 22; DB 1; Length 265;  
Best Local Similarity 100.0%; Pred. No. 1.7e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EQLGPVTQEFWDNLEKETEGLR 107  
|||||  
Db 85 EQLGPVTQEFWDNLEKETEGLR 106

RESULT 6

LPRB1Z

apolipoprotein A-I precursor (clone 22ap A1) - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999

C;Accession: S06064

R;Paraskevopoulou, T.B.; Kritlis, A.; Zannis, V.

submitted to the EMBL Data Library, July 1989

A;Reference number: S06064

A;Accession: S06064

A;Molecule type: mRNA

A;Residues: 1-266 <PAR>

A;Cross-references: EMBL:X15908; NID:g1457; PIDN:CAA34024.1; PID:g1458

C;Comment: This protein is synthesized in the small intestine.

C;Comment: This protein is a major component of the high density lipoproteins in plas

C;Superfamily: apolipoprotein A-I

C;Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-24/Domain: propeptide #status predicted <PRO>

F;25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 8.2%; Score 22; DB 1; Length 266;  
Best Local Similarity 100.0%; Pred. No. 1.7e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EQLGPVTQEFWDNLEKETEGLR 107  
|||||  
Db 85 EQLGPVTQEFWDNLEKETEGLR 106

RESULT 7

LPDGA1

apolipoprotein A-I precursor - dog

C;Species: Canis lupus familiaris (dog)

C;Date: 17-Dec-1982 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C;Accession: A60940; A03092; A61418

R;Luo, C.C.; Li, W.H.; Chan, L.

J. Lipid Res. 30, 1735-1746, 1989

A;Title: Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implic

A;Reference number: A60940; MUID:90132271

A;Accession: A60940

A;Molecule type: mRNA

A;Residues: 1-266 <LUO>

R;Chung, H.; Randolph, A.; Reardon, I.; Heinrichson, R.L.

J. Biol. Chem. 257, 2961-2967, 1982

A;Title: The covalent structure of apolipoprotein A-I from canine high density lipopr

A;Reference number: A03092; MUID:82142425

A;Accession: A03092

A;Molecule type: protein

A;Residues: 25-167, 'G', 169-201, 'Q', 203-234, 'Q', 236-266 <CHU>

R;Nakai, T.; Whayne, T.F.; Tang, J.

FEBS Lett. 64, 409-411, 1976

A;Title: The amino- and carboxyl-terminal sequences of canine apolipoprotein A-I.

A;Reference number: A61418; MUID:76210910

A;Accession: A61418

A;Molecule type: protein

A;Residues: 25-56, 'Z', 261-262, 'A' <NAK>

C;Superfamily: apolipoprotein A-I

C;Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid transport;

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-24/Domain: propeptide #status predicted <PRO>

F;25-266/Product: apolipoprotein A-I #status experimental <MAT>



Query Match 8.2%; Score 22; DB 1; Length 266;  
 A:Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolu  
 A:Reference number: I45853; MUID:90348478  
 A:Accession: I45853  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-265 <OXH>

QY 6 LTLAVLFLTGSQARHFWQDD 27  
 |||||  
 Db 6 LTLAVLFLTGSQARHFWQDD 27

RESULT 8  
 S31394

apolipoprotein A-I - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
 C:Accession: S31394  
 R:Moehel, B.; Flach, R.; Weiss, B.; Weller-Guettler, H.; Frey, A.; Zinke, H.; Gassen, H.  
 submitted to the EMBL Data Library, November 1992  
 A:Description: Genomic organization of the porcine apolipoprotein AI gene and study of g  
 A:Reference number: S31394  
 A:Accession: S31394  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-264 <MOE>  
 A:Cross-references: EMBL:X69477; NID:g1889; PIDN:CAA49234.1; PID:g1890  
 C:Superfamily: apolipoprotein A-I

Query Match 7.9%; Score 21; DB 2; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLTAVLFLTGSQARHFWQDD 25  
 |||||  
 Db 5 VLTAVLFLTGSQARHFWQDD 25

RESULT 9  
 A46018

apolipoprotein AI - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
 C:Accession: A46018  
 R:Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.  
 Genomics 15, 643-652, 1993  
 A:Title: Characterization of the apolipoprotein AI and CIII genes in the domestic pig.  
 A:Reference number: A46018; MUID:93224154  
 A:Accession: A46018  
 A:Status: preliminary  
 A:Molecule type: DNA; protein  
 A:Residues: 1-265 <BIR>  
 A:Cross-references: GB:I00626; NID:g164358; PIDN:AAA30992.1; PID:g164359  
 A:Note: sequence extracted from NCBI backbone (NCBIN:129509, NCBIP:129511)  
 C:Superfamily: apolipoprotein A-I

Query Match 7.9%; Score 21; DB 2; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLTAVLFLTGSQARHFWQDD 25  
 |||||  
 Db 5 VLTAVLFLTGSQARHFWQDD 25

RESULT 10

A56858  
 apolipoprotein A-I precursor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 05-Jan-1996 #sequence\_revision 23-Aug-1997 #text\_change 13-Aug-1999  
 C:Accession: I45853; A56858; A34649  
 R:O'Huigin, C.; Chan, L.; Li, W.  
 Mol. Biol. Evol. 7, 327-339, 1990

A:Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolu  
 A:Reference number: I45853; MUID:90348478  
 A:Accession: I45853  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-265 <OXH>

A:Cross-references: GB:M35870; NID:g162677; PIDN:AAA30381.1; PID:g162678  
 R:Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Auboliron, S.; Bauchart, D.; Chapman, M.J.;  
 Biochim. Biophys. Acta 1123, 145-150, 1992  
 A:Title: Plasma lipid transport in the pre-ruminant calf, Bos spp: primary structure o  
 A:Reference number: A56858; MUID:92153895  
 A:Accession: A56858  
 A:Status: preliminary

A:Molecule type: protein  
 A:Residues: 19-184, 'QL', 187-265 <SPA>

A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
 A:Note: sequence extracted from NCBI backbone (NCBIP:83520)

R:Auboliron, S.; Sparrow, D.A.; Beaubatie, L.; Bauchart, D.; Sparrow, J.T.; Laplaud, P  
 Biochem. Biophys. Res. Commun. 166, 833-839, 1990

A:Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasm  
 A:Reference number: A34649; MUID:90147795  
 A:Accession: A34649

A:Molecule type: protein  
 A:Residues: 25-70 <AUB>

A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks  
 C:Superfamily: apolipoprotein A-I

C:Keywords: lipid binding; lipoprotein  
 F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 7.9%; Score 21; DB 2; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLTAVLFLTGSQARHFWQDD 25  
 |||||  
 Db 5 VLTAVLFLTGSQARHFWQDD 25

RESULT 11  
 JT0672

apolipoprotein A-I - pig  
 N:Alternate names: apo-A-I  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 11-May-2000  
 C:Accession: JT0672; PN0471; A05311  
 R:Trieu, V.N.; Patel, B.; Zhan, R.; Black, D.D.  
 Gene 134, 267-270, 1993  
 A:Title: Sequence of the porcine apoA-I gene.  
 A:Reference number: JT0672; MUID:94085789  
 A:Accession: JT0672  
 A:Molecule type: DNA  
 A:Residues: 1-265 <TRI>  
 A:Cross-references: EMBL:Z14124; NID:g1893  
 A:Note: this translation is not annotated in GenBank entry SSAPOAIG, release 111.0; t  
 R:Trieu, V.N.; Hasler-Rapacz, J.; Rapacz, J.; Black, D.D.  
 Gene 123, 173-179, 1993  
 A:Title: Sequences and expression of the porcine apolipoprotein A-I and C-III mRNAs.  
 A:Reference number: PN0471; MUID:93154581  
 A:Accession: PN0471  
 A:Molecule type: mRNA  
 A:Residues: 105-265 <TR2>

A:Experimental source: liver  
 R:Mahley, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.  
 Biochemistry 15, 1928-1933, 1976

A:Title: Characterization of the plasma lipoproteins and apoproteins of the Erythroce  
 A:Reference number: A90395; MUID:76184721  
 A:Accession: A05311

A:Molecule type: protein

A:Residues: 25-34 <MAH>

C:Comment: This protein is the major apolipoprotein of high-density lipoprotein and s  
 C:Genetics:



A:Gene: apoA-I  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid metabolism; lip  
F:99/Region: ochre stop codon

Query Match 7.9%; Score 21; DB 2; Length 265;  
Best Local Similarity 100.0%; Pred. No. 1.6e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VLTAVLFLTGSGARHFWQD 25  
|||||  
DB 5 VLTAVLFLTGSGARHFWQD 25

RESULT 12  
JQ0704

apolipoprotein A-I - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 28-Oct-1994  
C:Accession: JQ0704  
R:Weiler-Guetli, H.; Sommerfeldt, M.; Papandriopoulou, A.; Mischek, U.; Bonitz, D.; F  
J. Neurochem. 54, 444-450, 1990  
A:Title: Synthesis of apolipoprotein A-I in pig brain microvascular endothelial cells.  
A:Reference number: JQ0704; MUID:90132667  
A:Accession: JQ0704  
A:Molecule type: mRNA  
A:Residues: 1-231 <MEI>  
A:Note: the authors translated the codon CAG for residue 124 as His and GAC for residue  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; lipid binding; lipid transport; lipoprotein; pl

Query Match 7.5%; Score 20; DB 2; Length 231;  
Best Local Similarity 100.0%; Pred. No. 1.4e-11;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 REQLGPVTOEFWDNLEKETE 104  
|||||  
DB 51 REQLGPVTOEFWDNLEKETE 70

RESULT 13  
S21830

apolipoprotein A-I - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 20-Feb-1995 #sequence\_revision 22-May-1998 #text\_change 13-Aug-1999  
C:Accession: S21830  
R:Trieu, V.N.; Rohwer-Nuttler, P.; Black, D.D.  
submitted to the EMBL Data Library, May 1991  
A:Description: Sequence and developmental expression of porcine ApoA-I mRNA.  
A:Reference number: S21830  
A:Accession: S21830  
A:Molecule type: mRNA  
A:Residues: 1-164 <TRI>  
A:Cross-references: EMBL:X59414; NID:g1891; PIDN:CAA42050.1; PID:g1892  
A:Experimental source: liver  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid metabolism; lip

Query Match 7.1%; Score 19; DB 2; Length 164;  
Best Local Similarity 100.0%; Pred. No. 9.5e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 229 EKAKPALEDLRQGLLPVLE 247  
|||||  
DB 126 EKAKPALEDLRQGLLPVLE 144

RESULT 14  
A24700

apolipoprotein A-I precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 13-Aug-1999  
C:Accession: A24700; S00298; A05314  
R:Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K.  
J. Biol. Chem. 261, 13268-13277, 1986  
A:Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and  
A:Reference number: A92558; MUID:87008540  
A:Accession: A24700

A:Molecule type: DNA  
A:Residues: 1-259 <HAD>  
A:Cross-references: EMBL:J02597; NID:g202935; PIDN:AAA40745.1; PID:g202939  
R:Poncin, J.E.; Martial, J.A.; Gjelten, J.E.  
Eur. J. Biochem. 140, 493-498, 1984  
A:Title: Cloning and structure analysis of the rat apolipoprotein A-I cDNA.  
A:Reference number: S00298; MUID:84207987  
A:Accession: S00298

A:Molecule type: mRNA  
A:Residues: 1-259 <PON>  
A:Cross-references: GB:M00001; EMBL:X00558; NID:g202944; PIDN:AAA40749.1; PID:g202945  
R:Gordon, J.I.; Smith, D.P.; Andy, R.; Alpers, D.H.; Schonfeld, G.; Strauss, A.W.  
J. Biol. Chem. 257, 971-978, 1982  
A:Title: The primary translation product of rat intestinal apolipoprotein A-I mRNA is  
A:Reference number: A05314; MUID:82098162  
A:Accession: A05314

A:Molecule type: protein  
A:Residues: 1-18, 'X', 20-21, 'X', 23-28, 'X', 30-38, 'X', 40-41, 'X', 43-45 <GOR>  
C:Comment: This protein is synthesized in the liver and small intestine. The propepti  
C:Genetics: This protein is a major component of the high density lipoproteins in plas  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; l  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-259/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 4.5%; Score 12; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 KDSGRDYVSQFE 58  
|||||  
DB 46 KDSGRDYVSQFE 57

RESULT 15  
JC1237

apolipoprotein A-I precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-1993  
C:Accession: JC1237  
R:Boyle, T.P.; Marotti, K.R.  
Gene 117, 243-247, 1992  
A:Title: Structure of the murine gene encoding apolipoprotein A-I.  
A:Reference number: JC1237; MUID:92347700  
A:Accession: JC1237  
A:Molecule type: DNA  
A:Residues: 1-262 <BOY>  
A:Cross-references: GB:M77801  
C:Genetics:

A:Gene: ApoA-I  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; l  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-262/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 4.5%; Score 12; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 KDSGRDYSQFE 58  
|||||  
Db 46 KDSGRDYSQFE 57

## RESULT 16

S22420

apolipoprotein A-I precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
C:Accession: S22420; S22421; A44364  
R:Stoffel, W.; Mueller, R.; Binczek, E.; Hofmann, K.  
Biol. Chem. Hoppe-Seyler 373, 187-193, 1992  
A:Title: Mouse apolipoprotein AI. cDNA-derived primary structure, gene organisation and  
A:Reference number: S22420; MUID:92281682  
A:Accession: S22420  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-264 <STO>  
A:Cross-references: EMBL:X64262; NID:g50014; PIDN:CAA45560.1; PID:g50015  
A:Accession: S22421  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <ST2>  
A:Cross-references: EMBL:X64263; NID:g50020; PIDN:CAA45561.1; PID:g50021  
R:Janusz, J.L.; Azrolan, N.; O'Connell, A.; Aalto-Setälä, K.; Breslow, J.L.  
Genomics 14, 1081-1088, 1992  
A:Title: Characterization of the mouse apolipoprotein ApoA-1/Apoc-3 gene locus: genomic,  
A:Reference number: A44364; MUID:93122774  
A:Accession: A44364  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-264 <JAN>  
A:Note: sequence extracted from NCBI backbone (NCBIN:122400, NCBI:P:122407)  
C:Genetics:  
A:Introns: 15/1; 66/2  
C:Superfamily: apolipoprotein A-I

Query Match 4.5%; Score 12; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 KDSGRDYSQFE 58  
|||||  
Db 46 KDSGRDYSQFE 57

## RESULT 17

G83132

hypothetical protein PA4111 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: G83132  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A:Reference number: A82950; MUID:20437337  
A:Accession: G83132  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-132 <STO>  
A:Cross-references: GB:AE004827; GB:AE004091; NID:g9950306; PIDN:AAG07498.1; GSPDB:GN001  
C:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA4111

Query Match 3.7%; Score 10; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 0.052;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 196 ORLARLEAL 205  
|||||  
Db 41 ORLARLEAL 50

## RESULT 18

A61448

apolipoprotein A-I - duck  
C:Species: Anas platyrhynchos (domestic duck)  
C:Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 17-Mar-1999  
C:Accession: A61448  
R:Gu, Z.W.; Xie, Y.H.; Yang, M.; Sparrow, J.T.; Wang, K.; Li, Y.; Li, W.H.; Gotto Jr.  
J. Protein Chem. 12, 585-591, 1993  
A:Title: Primary structure of Beijing duck apolipoprotein A-1.  
A:Reference number: A61448; MUID:94190456  
A:Accession: A61448  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-246 <GUA>  
A:Note: the source is designated as Beijing duck  
C:Superfamily: apolipoprotein A-I

Query Match 3.7%; Score 10; DB 2; Length 246;  
Best Local Similarity 100.0%; Pred. No. 0.093;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 187 LAPYSDELRLQ 196  
|||||  
Db 168 LAPYSDELRLQ 177

## RESULT 19

LPCHA1

apolipoprotein A-I precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Mar-1990 #sequence\_revision 25-Aug-1995 #text\_change 22-Jun-1999  
C:Accession: JH0471; A29657; S01453; S28888; A29616; S00187; S10973; S11023; I50156  
R:Bhattacharyya, N.; Chattapadhyay, R.; Hirsch, A.; Banerjee, D.  
Gene 104, 163-168, 1991  
A:Title: Isolation, characterization and sequencing of the chicken apolipoprotein-AI-  
A:Reference number: JH0471; MUID:92009209  
A:Accession: JH0471  
A:Molecule type: DNA  
A:Residues: 1-264 <BHA>  
R:Byrnes, L.; Luo, C.C.; Li, W.H.; Yang, C.; Chan, L.  
Biochem. Biophys. Res. Commun. 148, 485-492, 1987  
A:Title: Chicken apolipoprotein A-I: cDNA sequence, tissue expression and evolution.  
A:Reference number: A29657; MUID:88049703  
A:Accession: A29657  
A:Molecule type: mRNA  
A:Residues: 1-264 <BYR>  
A:Cross-references: EMBL:M17961; NID:g211147; PIDN:AAA48593.1; PID:g211148  
R:Rajavashisth, T.B.; Dawson, P.A.; Williams, D.L.; Shackelford, J.E.; Leberer, H.; L  
J. Biol. Chem. 262, 7058-7065, 1987  
A:Title: Structure, evolution, and regulation of chicken apolipoprotein A-I.  
A:Reference number: S01453; MUID:8722301  
A:Accession: S01453  
A:Molecule type: mRNA  
A:Residues: 1-15, 'I', 17-147, 'K', 149-264 <RAJ>  
A:Cross-references: GB:M25559; EMBL:J02739; NID:g211145; PIDN:AAA48592.1; PID:g211146  
A:Accession: S28888  
A:Molecule type: protein  
A:Residues: 25-44; 230-256 <RA2>  
R:Ferrari, S.; Tarugli, P.; Drusiani, E.; Calandra, S.; Fregni, M.  
Gene 60, 39-46, 1987  
A:Title: The complete sequence of chick apolipoprotein AI mRNA and its expression in  
A:Reference number: A29616; MUID:88152500  
A:Accession: A29616  
A:Molecule type: mRNA  
A:Residues: 1-264 <FER>

A;Cross-references: EMBL:M18746; NID:g211149; PIDN:AAA48594.1; PID:g211150  
R;Yang, C.Y.; Gu, Z.W.; Patsch, W.; Weng, S.A.; Kim, T.W.; Chan, L.  
FEBS Lett. 224, 261-266, 1987  
A;Title: The complete amino acid sequence of proapolipoprotein A-I of chicken high density lipoprotein  
A;Reference number: S00187; MUID:88083548  
A;Accession: S00187  
A;Molecule type: protein  
A;Residues: 19-264 <YAN>  
R;Banerjee, D.; Mukherjee, T.K.; Redman, C.M.  
J. Cell Biol. 101, 1219-1226, 1985  
A;Title: Biosynthesis of high density lipoprotein by chicken liver: intracellular transport and secretion  
A;Reference number: S10973; MUID:86008443  
A;Accession: S10973  
A;Molecule type: protein  
A;Residues: 19, 'H', 21-23, 'Q', 25-44 <BAN>  
A;Note: the sequence from Fig. 4 is inconsistent with that from Table II in having 41-Threonine at position 25  
R;Shackelford, J.E.; Leberer, H.G.  
J. Biol. Chem. 258, 7175-7180, 1983  
A;Title: Synthesis and secretion of apolipoprotein A1 by chick breast muscle.  
A;Reference number: S11023; MUID:83213468  
A;Accession: S11023  
A;Molecule type: protein  
A;Residues: 25-44 <SHA>  
R;Lamon-Fava, S.; Sastry, R.; Ferrarri, S.; Rajavashisth, T.B.; Lusis, A.J.; Karathanasis, S.K.  
J. Lipid Res. 33, 831-842, 1992  
A;Title: Evolutionary distinct mechanisms regulate apolipoprotein A-I gene expression: Divergent regulation of the A-I and A-II genes  
A;Reference number: I50156; MUID:92381402  
A;Accession: I50156  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-15, 'I', 17-264 <LAM>  
A;Cross-references: GB:M96012; NID:g211158; PIDN:AAA48597.1; PID:g211159  
C;Comment: This protein is synthesized only in the liver in mammals, whereas in chicken it is synthesized in the liver and in the yolk of the egg  
C;Comment: This protein is a major component of the high density lipoproteins in plasma.  
C;Genetics:  
A;Gene: ApoA1  
A;Introns: 15/1; 66/2  
C;Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lipoprotein; signal sequence  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-264/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 3.7%; Score 10; DB 1; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.099;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 LAPYSDLRQ 196  
| | | | | | | | | |  
Db 186 LAPYSDLRQ 195

RESULT 20  
apolipoprotein A-I precursor - Japanese quail  
C;Species: Coturnix coturnix japonica (Japanese quail)  
C;Date: 02-Sep-1997 #sequence\_revision 05-Sep-1997 #text\_change 20-Jun-2000  
C;Accession: JC5456; PC4472  
R;Oku, H.; Toda, T.; Nagata, J.; Ishikawa, M.; Neyazaki, K.; Shinjo, C.; Chinen, I.  
Biochem. Biotechnol. Biochem. 61, 286-290, 1997  
A;Title: Apolipoprotein A-I of Japanese quail: cDNA sequence and modulation of tissue expression  
A;Reference number: JC5456; MUID:97212027  
A;Accession: JC5456  
A;Molecule type: mRNA  
A;Residues: 1-264 <OKU>  
A;Cross-references: GB:D85133; NID:g1890065; PIDN:BAAL2729.1; PID:g1890066  
A;Accession: PC4472  
A;Molecule type: protein  
A;Residues: 25-60 <OK2>  
C;Comment: This protein is the major protein moiety of plasma HPL. It functions as a very low density lipoprotein  
F;1-24/Domain: signal sequence #status predicted <SIG>

F;25-264/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 3.4%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.95;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLAVFLTG 15  
| | | | | | | | | |  
Db 7 TLAVFLTG 15

RESULT 21  
A05313  
apolipoprotein A-I - red guenon (fragment)  
N;Alternate names: apo-A-I  
C;Species: Erythrocebus patas (red guenon, hussar)  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 11-May-2000  
C;Accession: A05313  
R;Mahley, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.  
Biochemistry 15, 1928-1933, 1976  
A;Title: Characterization of the plasma lipoproteins and apoproteins of the Erythrocebus patas  
A;Reference number: A90395; MUID:76184721  
A;Accession: A05313  
A;Molecule type: protein  
A;Residues: 1-20 <MAH>  
C;Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; 1

Query Match 3.0%; Score 8; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 PWDVKDL 38  
| | | | | | | | | |  
Db 7 PWDVKDL 14

RESULT 22  
G83100  
hypothetical protein PA4368 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: G83100  
R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lior, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A;Reference number: A82950; MUID:20437337  
A;Accession: G83100  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-300 <STO>  
A;Cross-references: GB:AE004852; GB:AE004091; NID:g9950587; PIDN:AAG07756.1; GSPDB:GN000001  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA4368

Query Match 3.0%; Score 8; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 RLARLEA 204  
| | | | | | | | | |  
Db 228 RLARLEA 235

RESULT 23  
S33450  
apolipoprotein E - pig

N;Alternate names: apo-E  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
C;Accession: S33450; A05312  
R;Brzozowska, A.M.; Grimholt, U.; Kulseth, M.A.; Wold, I.; Rogne, S.  
submitted to the EMBL Data Library, May 1993  
A;Description: The sequence of porcine apolipoprotein E(APOE) cDNA.  
A;Reference number: S33450  
A;Accession: S33450  
A;Molecule type: mRNA  
A;Residues: 1-317 <BR2>  
A;Cross-references: EMBL:X72835; NID:g311232; PIDN:CAA51356.1; PID:g311233  
R;Weisgraber, K.H.; Troxler, R.F.; Rall, S.C.; Mahley, R.W.  
Biochem. Biophys. Res. Commun. 95, 374-380, 1980  
A;Reference number: A90221; MUID:81021043  
A;Accession: A05312  
A;Molecule type: protein  
A;Residues: 19-26,'Q',28,'GXL' <WEI>  
C;Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol; chylomicron; HDL; lipid binding; lipid transport; lipoprotein;

Query Match 3.0%; Score 8; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 87 QLGPTQ 94  
|||||  
Db 98 QLGPTQ 105

## RESULT 24

F82692  
conserved hypothetical protein XF1359 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: F82692  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: F82692  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-387 <STIM>  
A;Cross-references: GB:AE003967; GB:AE003849; NID:g9106347; PIDN:AAF84168.1; GSPDB:GN001  
A;Experimental source: strain 9a5c  
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briñones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Veltore, A.L.; Z  
A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF1359

Query Match 3.0%; Score 8; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 AAVLTAV 10  
|||||  
Db 161 AAVLTAV 168

RESULT 25  
B40892  
apolipoprotein A-IV precursor - mouse (strain 129)  
C;Species: Mus musculus (house mouse)  
C;Date: 27-Mar-1992 #sequence\_revision 10-Apr-1992 #text\_change 13-Aug-1999  
C;Accession: B40892  
R;Reue, K.; Leete, T.H.  
J. Biol. Chem. 266, 12715-12721, 1991  
A;Title: Genetic variation in mouse apolipoprotein A-IV due to insertion and deletion  
A;Reference number: A40892; MUID:91286309  
A;Accession: B40892  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-391 <REU>  
A;Cross-references: GB:M64248; NID:g191884; PIDN:AAA37214.1; PID:g191885  
A;Note: the authors translated the codon CTG for residue 87 as Glu, GAG for residue 2  
C;Superfamily: apolipoprotein A-I  
C;Keywords: lipid binding; lipoprotein

Query Match 3.0%; Score 8; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KAAVLT 9  
|||||  
Db 4 KAAVLT 11

## RESULT 26

A25281  
apolipoprotein A-IV precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 13-Aug-1999  
C;Accession: A25281  
R;Williams, S.C.; Bruckheimer, S.M.; Lusis, A.J.; LeBoeuf, R.C.; Kinniburgh, A.J.  
Mol. Cell. Biol. 6, 3807-3814, 1986  
A;Title: Mouse apolipoprotein A-IV gene: nucleotide sequence and induction by a high-  
A;Reference number: A25281; MUID:87089722  
A;Accession: A25281  
A;Molecule type: DNA  
A;Residues: 1-394 <WIL>  
A;Cross-references: GB:M13966; NID:g192006; PIDN:AAA37253.1; PID:g387102  
C;Genetics:  
A;Introns: 16/1; 58/2  
C;Superfamily: apolipoprotein A-I  
C;Keywords: chylomicron; HDL; lipid binding; lipoprotein

Query Match 3.0%; Score 8; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KAAVLT 9  
|||||  
Db 4 KAAVLT 11

## RESULT 27

A40892  
apolipoprotein A-IV precursor - mouse (strain C57BL16)  
C;Species: Mus musculus (house mouse)  
C;Date: 27-Mar-1992 #sequence\_revision 10-Apr-1992 #text\_change 13-Aug-1999  
C;Accession: A40892  
R;Reue, K.; Leete, T.H.  
J. Biol. Chem. 266, 12715-12721, 1991  
A;Title: Genetic variation in mouse apolipoprotein A-IV due to insertion and deletion  
A;Reference number: A40892; MUID:91286309  
A;Accession: A40892  
A;Status: preliminary  
A;Molecule type: mRNA



A;Residues: 1-395 <REU>  
A;Cross-references: GB:M64249; NID:g191886; PIDN:AAA37215.1; PID:g191887  
A;Note: the authors translated the codon CTG for residue 87 as Glu, GAG for residue 207  
C;Superfamily: apolipoprotein A-I  
C;Keywords: lipid binding; lipoprotein

Query Match 3.0%; Score 8; DB 2; Length 395;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAAVLTIA 9  
|||||  
Db 4 KAAVLTIA 11

RESULT 28

C40892 apolipoprotein A-IV precursor - mouse

C;Species: Mus musculus castaneus (southeastern Asian house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 31-Dec-1993

C;Accession: C40892

R;Reue, K.; Leete, T. H.

J. Biol. Chem. 266, 12715-12721, 1991

A;Title: Genetic variation in mouse apolipoprotein A-IV due to insertion and deletion in  
A;Reference number: A40892; MUID:91286309

A;Accession: C40892

A;Molecule type: mRNA

A;Residues: 1-399 <REU>

A;Cross-references: GB:M64250

A;Note: the authors translated the codon CTG for residue 87 as Glu, GAG for residue 207

C;Superfamily: apolipoprotein A-I

C;Keywords: chylomicron; duplication; HDL; lipid binding; lipid transport; lipoprotein;

Query Match 3.0%; Score 8; DB 2; Length 399;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAAVLTIA 9  
|||||  
Db 4 KAAVLTIA 11

RESULT 29

S12332 ubiquitin--protein ligase (EC 6.3.2.19) - yeast (Saccharomyces cerevisiae) (strain S288C)

N;Alternate names: protein G7168; protein YGR184C

C;Species: Saccharomyces cerevisiae

C;Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 20-Jun-2000

C;Accession: S12332; S64502; S64498

R;Barfel, B.; Wuening, I.; Varshavsky, A.

EMBO J. 9, 3179-3189, 1990

A;Title: The recognition component of the N-end rule pathway.

A;Reference number: S12332; MUID:91006011

A;Accession: S12332

A;Molecule type: DNA

A;Residues: 1-1950 <BAR>

A;Cross-references: EMBL:X53747; NID:g4743; PIDN:CAA37779.1; PID:g4744

A;Experimental source: strain S288C

R;Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M. I.; Sanchez-Perez, M.; Nombela, C.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S64499

A;Accession: S64502

A;Molecule type: DNA

A;Residues: 1-1950 <ARR>

A;Cross-references: EMBL:Z72969; NID:g132325; PIDN:CAA97210.1; PID:g132326; MIPS:YGR18

A;Experimental source: strain S288C

R;Hebling, U.; Hofmann, B.; Delius, H.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S64003

A;Accession: S64498

A;Molecule type: DNA

A;Residues: 1615-1950 <HEB>  
A;Cross-references: EMBL:Z72969; MIPS:YGR184C

A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:UBR1; PTR1

A;Cross-references: SGD:S0003416; MIPS:YGR184C

A;Map position: 7R

C;Superfamily: ubiquitin--protein ligase

C;Keywords: ligase; protein degradation

Query Match 3.0%; Score 8; DB 2; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 LEALKENG 209  
|||||  
Db 1277 LEALKENG 1284

RESULT 30

T05224 hypothetical protein F1715.150 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999

C;Accession: T05224

R;Bevan, M.; Vitale, D.; Liguori, R.; Argirion, A.; De Simone, V.; Hoheisel, J.; Mewe

submitted to the Protein Sequence Database, July 1998

A;Accession: T05224

A;Molecule type: DNA

A;Residues: 1-67 <BEV>

A;Cross-references: EMBL:AL031032

A;Experimental source: cultivar Columbia; BAC clone F1715

C;Genetics:

A;Map position: 4

A;Note: F1715.150

Query Match 2.6%; Score 7; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 238 LRQGLLP 244  
|||||  
Db 46 LRQGLLP 52

RESULT 31

H87207 probable membrane protein [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C;Accession: H87207

R;Cole, S. T.; Elgimeier, K.; Parkhill, J.; James, K. D.; Thomson, N. R.; Wheeler, P. R.;

R.; Davies, R. M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holro

eam, M. A.; Rutherford, K. M.

Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;

A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21128732; PMID:11234002

A;Accession: H87207

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-101 <STO>

A;Cross-references: GB:AL450380; NID:g13093975; PIDN:CAC31904.1; GSPDB:GN00147

C;Genetics:

A;Gene: ML2388

Query Match 2.6%; Score 7; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 5 VLTLAVL 11  
|||||  
Db 90 VLTLAVL 96

## RESULT 32

E71566

probable S16 ribosomal protein - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999

C:Accession: E71566

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac

A:Reference number: A71570; MUID:99000809

A:Accession: E71566

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 &lt;ARN&gt;

A:Cross-references: GB:AE001277; GB:AE001273; NID:g3328408; PIDN:AAC67616.1; PID:g332841

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: rsl6

C:Superfamily: Escherichia coli ribosomal protein S16

## Query Match

Best Local Similarity 2.6%; Score 7; DB 2; Length 116;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 RQRLAAR 201  
|||||  
Db 103 RQRLAAR 109

## RESULT 33

C86709

hypothetical protein yghC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C:Accession: C86709

R:Boletín, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: C86709

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-132 &lt;STO&gt;

A:Cross-references: GB:AE005176; PID:g12723583; PIDN:AAK04773.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yghC

Query Match 2.6%; Score 7; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 45;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 207 ENGGAARL 213  
|||||  
Db 122 ENGGAARL 128

## RESULT 34

G72710

hypothetical protein APE1102 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C:Accession: G72710

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: G72710

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-139 &lt;KAW&gt;

A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA80087.1; PID:d1043873; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1102

## Query Match

Best Local Similarity 2.6%; Score 7; DB 2; Length 139;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 QRLAARL 202  
|||||  
Db 120 QRLAARL 126

## RESULT 35

T33928

hypothetical protein Y66H1A.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jun-2000

C:Accession: T33928

R:Clarke, K.; Wohldmann, P.

submitted to the EMBL Data Library, February 1999

A:Description: The sequence of C. elegans cosmid Y66H1A.

A:Reference number: Z21440

A:Accession: T33928

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-150 &lt;CLA&gt;

A:Cross-references: EMBL:AF125462; PIDN:AAD12855.1; GSPDB:GN00022; CESP:Y66H1A.3

A:Experimental source: strain Bristol N2; clone Y66H1A

C:Genetics:

A:Gene: CESP:Y66H1A.3

A:Map position: 4

A:Introns: 17/1; 50/1; 92/3; 124/3

C:Superfamily: Caenorhabditis elegans hypothetical protein Y66H1A.3

Query Match 2.6%; Score 7; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 51;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 RQRLAAR 201  
|||||  
Db 109 RQRLAAR 115

## RESULT 36

E86711

hypothetical protein y9jB [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C:Accession: E86711

R:Boletín, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehr Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: E86711

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-150 &lt;STO&gt;

A:Cross-references: GB:AE005176; PID:g12723602; PIDN:AAK04791.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: y9jB

Query Match 2.6%; Score 7; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 KPALEDL 238  
|||||  
Db 136 KPALEDL 142

## RESULT 37

AB0918

conserved hypothetical protein STY3603 [imported] - Salmonella enterica subsp. enterica

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001

C:Accession: AB0918

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AB0918

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD07936.1; PID:g16504481; GSPDB:GN00176

C:Genetics:

A:Gene: STY3603

Query Match 2.6%; Score 7; DB 2; Length 155;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 DELRQRL 198  
|||||  
Db 90 DELRQRL 96

## RESULT 38

F91222

hypothetical protein ECS4750 [imported] - Escherichia coli (strain O157:H7, substrain R1

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: F91222

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F91222

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA838173.1; PID:g13364226; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECS4750

Query Match 2.6%; Score 7; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 DELRQRL 198  
|||||  
Db 96 DELRQRL 102

## RESULT 39

## S30747

hypothetical 17.1K protein (rard-pldA intergenic region) - Escherichia coli

N:Alternate names: hypothetical protein fl61

C:Species: Escherichia coli

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999

C:Accession: S30747; E65186; S30710

R:Ohmori, H.

submitted to the EMBL Data Library, January 1993

A:Description: Physical map of the corA region of the E.coli chromosome.

A:Reference number: S30742

A:Accession: S30747

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <OHM>

A:Cross-references: EMBL:L02122; NID:g145575; PIDN:AAD15042.1; PID:g145581

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.: Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: E65186

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-161 <BLAT>

A:Cross-references: GB:AE000457; GB:U00096; NID:g2367294; PIDN:AACT6823.1; PID:g23672

A:Experimental source: strain K-12, substrain MGL655

R:Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.

Science 257, 771-778, 1992

A:Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.

A:Reference number: S30660; MUID:92358234

A:Accession: S30710

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-101, 'V', 103-161 <DAN>

A:Cross-references: EMBL:M87049; NID:g836656; PIDN:AAA67616.1; PID:g148219

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992

C:Genetics:

A:Gene: y1g1

Query Match 2.6%; Score 7; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 DELRQRL 198  
|||||  
Db 96 DELRQRL 102

## RESULT 40

D86069

hypothetical protein y1g1 [imported] - Escherichia coli (strain O157:H7, substrain ED

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: D86069

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; DiMalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D86069

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-161 <STO>

A:Cross-references: GB:AE005174; NID:g12518694; PIDN:AAG59016.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: y1g1

Query Match 2.6%; Score 7; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 192 DELRQRL 198  
|||||||  
Db 96 DELRQRL 102

## RESULT 41

H95871

conserved hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magapla  
C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C/Accession: H95871

R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A/Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo

A/Reference number: A95842; MUID:21396508; PMID:11481431

A/Accession: H95871

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-199 <KUR>

A/Cross-references: GB:AL591985; PIDN:CAC48640.1; PID:g15140112; GSPDB:GN00167

A/Experimental source: strain 1021, megaplasmid pSymb

R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Gene: Smb20250

A/Genome: plasmid

## Query Match

Best Local Similarity 2.6%; Score 7; DB 2; Length 199;  
Matches 7; Conservative 100.0%; Pred. No. 66;

Matches 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 QRLAARL 202

|||||||

Db 38 QRLAARL 44

## RESULT 42

H84357

hypothetical protein Vng2080c [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C/Accession: H84357

R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A/Title: Genome sequence of Halobacterium species NRC-1.

A/Reference number: A84160; MUID:20504483

A/Accession: H84357

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-205 <STO>

A/Cross-references: GB:AE004437; MID:g10581497; PIDN:AAG20228.1; GSPDB:GN00138

C/Genetics:

A/Gene: VNG2080C

## Query Match

Best Local Similarity 2.6%; Score 7; DB 2; Length 205;  
Matches 7; Conservative 100.0%; Pred. No. 68;

Matches 0; Mismatches 0; Indels 0; Gaps 0;

OY 177 RAHVDAAL 183

|||||||

Db 41 RAHVDAAL 47

## RESULT 43

T37221

hypothetical protein SC2H4.05 - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C/Accession: T37221

R/Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, September 1998

A/Reference number: Z21615

A/Accession: T37221

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-216 <OLI>

A/Cross-references: EMBL:AL031514; PIDN:CAA20599.1; GSPDB:GN00070; SCOEDB:SC2H4.05

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCOEDB:SC2H4.05

## Query Match

Best Local Similarity 2.6%; Score 7; DB 2; Length 216;  
Matches 7; Conservative 100.0%; Pred. No. 72;

Matches 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 GLLPVLE 247

|||||||

Db 89 GLLPVLE 95

## RESULT 44

E85021

hypothetical protein AT4g01670 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

C/Accession: E85021

R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999

A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A/Reference number: A85001; MUID:20083488

A/Accession: E85021

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-218 <STO>

A/Cross-references: GB:NC\_001268; MID:g7268210; PIDN:CAB77737.1; GSPDB:GN00140

C/Genetics:

A/Gene: AT4g01670

A/Map position: 4

C/Superfamily: Arabidopsis hypothetical protein F14M4.23

## Query Match

Best Local Similarity 2.6%; Score 7; DB 2; Length 218;  
Matches 7; Conservative 100.0%; Pred. No. 72;

Matches 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 DLEEVKA 119

|||||||

Db 172 DLEEVKA 178

## RESULT 45

G65089

Probable transcriptional regulatory protein b3025 - Escherichia coli (strain K-12)

C/Species: Escherichia coli

C/Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 27-Nov-2001

C/Accession: G65089

R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; MUID:97426617

A/Accession: G65089

A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-219 <BLAT>  
A;Cross-references: GB:AE000384; GB:U00096; NID:g2367186; PIDN:AAC76061.1; PID:g1789402;  
A;Experimental source: strain K-12, substrain MG1655  
C;Superfamily: ompR protein; response regulator homology  
C;Keywords: phosphoprotein  
F;3-112/Domain: response regulator homology <RRH>  
F;51/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 2.6%; Score 7; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205  
|||||  
Db 109 AARLEAL 115

## RESULT 46

C91117  
probable 2-component transcription regulator [Imported] - Escherichia coli (strain O157:

C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C;Accession: C91117

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinaagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C91117

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-219 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA837330.1; PID:g13363379; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RMD 0509952

C;Genetics:

A;Gene: ECS3907

C;Superfamily: ompR protein; response regulator homology

Query Match 2.6%; Score 7; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205  
|||||  
Db 109 AARLEAL 115

## RESULT 47

C85962  
probable 2-component transcription regulator ygiX [Imported] - Escherichia coli (strain

C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C;Accession: C85962

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: C85962

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-219 <STO>

A;Cross-references: GB:AE005174; NID:g12517595; PIDN:AAG58159.1; GSPDB:GN00145; UWGP:Z43

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: ygiX

C;Superfamily: ompR protein; response regulator homology

Query Match 2.6%; Score 7; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205  
|||||  
Db 109 AARLEAL 115

## RESULT 48

AI0888

probable two-component system response regulator. STY3354 [Imported] - Salmonella ent

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001

C;Accession: AI0888

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A;Reference number: AB0502; PMID:11677608

A;Accession: AI0888

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-219 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD03009.1; PID:g16504255; GSPDB:GN00176

C;Genetics:

A;Gene: STY3354

C;Superfamily: ompR protein; response regulator homology

Query Match 2.6%; Score 7; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205  
|||||  
Db 109 AARLEAL 115

## RESULT 49

F65018  
hypothetical protein b2439 - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 04-Mar-2000

C;Accession: F65018

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617

A;Accession: F65018

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-219 <BLAT>

A;Cross-references: GB:AE000331; GB:U00096; NID:g1788775; PIDN:AAC75492.1; PID:g17887

A;Experimental source: strain K-12, substrain MG1655

C;Superfamily: Escherichia coli hypothetical protein b2439

Query Match 2.6%; Score 7; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FLTGSQA 18  
|||||  
Db 189 FLTGSQA 195

## RESULT 50

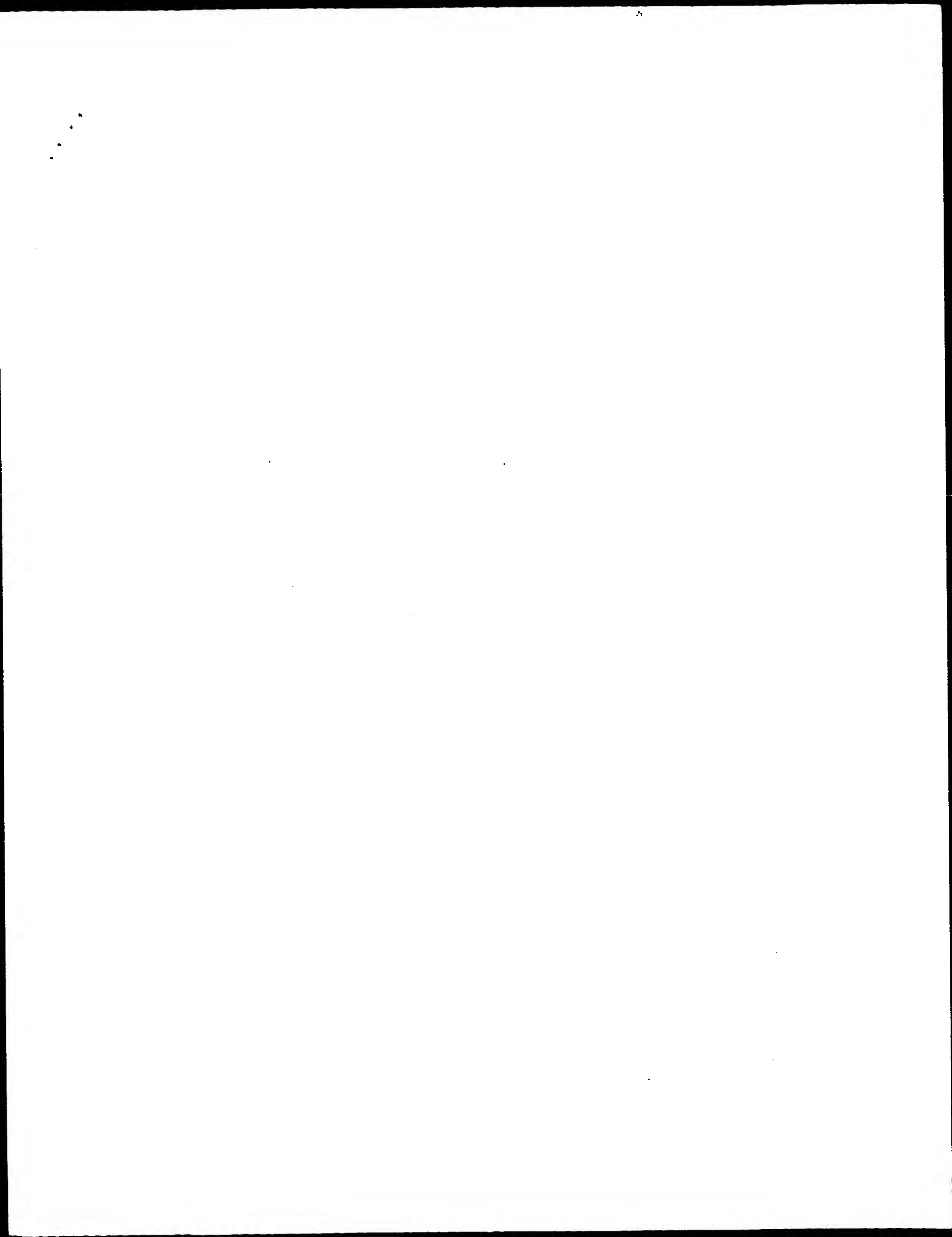
F91042

hypothetical protein ECS3310 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
 C;Species: Escherichia coli  
 C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C;Accession: F91042  
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
 A;Reference number: A99629; MUID:21156231; PMID:11258796  
 A;Accession: F91042  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-219 <HAY>  
 A;Cross-references: GB:BA000007; PIDN:BAB36733.1; PID:g13362780; GSPDB:GN00154  
 A;Experimental source: strain O157:H7, substrain R1MD 0509952  
 C;Genetics:  
 A;Gene: ECS3310  
 C;Superfamily: Escherichia coli hypothetical protein b2439

Query Match 2.6%; Score 7; DB 2; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 FLTGSQA 18  
 |||||  
 Db 189 FLTGSQA 195

Search completed: September 22, 2002, 12:23:38  
 Job time: 233 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:23:20 ; Search time 11.64 Seconds

(without alignments)  
888.155 Million cell updates/sec

Title: US-09-803-918A-2

Perfect score: 267  
Sequence: 1 MKAAYLTLAVFLTGSQARH.....SFKVSFLSALFEYTKKLNTO 267

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	100.0	267	1	APAL_HUMAN
2	69	25.8	267	1	APAL_MACFA
3	23	8.6	265	1	APAL_TJUGB
4	22	8.2	266	1	APAL_CANFA
5	22	8.2	266	1	APAL_RABIT
6	21	7.9	265	1	APAL_BOVIN
7	21	7.9	265	1	APAL_PIG
8	12	4.5	259	1	APAL_PIG
9	12	4.5	264	1	APAL_MOUSE
10	10	3.7	264	1	APAL_ANAPL
11	10	3.7	264	1	APAL_CHICK
12	9	3.4	264	1	APAL_COTJA
13	8	3.0	20	1	APAL_ERIPA
14	8	3.0	317	1	APE_PIG
15	8	3.0	395	1	APAL_MOUSE
16	8	3.0	1950	1	UBRI_YEAST
17	7	2.6	116	1	RS16_CHLMU
18	7	2.6	116	1	RS16_CHLTR
19	7	2.6	132	1	TRMA_LACLA
20	7	2.6	139	1	YB02_AERPE
21	7	2.6	155	1	YIGI_ECOLI
22	7	2.6	155	1	YIGI_ECOLI
23	7	2.6	155	1	YIGI_SALTY
24	7	2.6	207	1	LOLB_ACTAC
25	7	2.6	219	1	EUTL_ECOLI
26	7	2.6	219	1	EUTL_SALTY
27	7	2.6	219	1	YGTX_ECOLI
28	7	2.6	250	1	CHIT_DIOJA
29	7	2.6	273	1	SC65_YEAST
30	7	2.6	300	1	VAT1_HUMAN
31	7	2.6	314	1	RPOA_BACSU
32	7	2.6	316	1	Y034_TREPA
33	7	2.6	330	1	EXI2_HUMAN
			349	1	DIA2_MOUSE

34	7	2.6	352	1	REP3_THIFE	P27186 thlobacillu
35	7	2.6	359	1	C1ID_MOUSE	Q96d19 mus musculu
36	7	2.6	366	1	P53_PLAIE	O12946 platichthys
37	7	2.6	379	1	YCC2_ECOLI	P75881 escherichia
38	7	2.6	394	1	DXR_NEIMA	O9jx33 neisseria m
39	7	2.6	394	1	DXR_NEIMA	O9k198 neisseria m
40	7	2.6	394	1	HYF1_ALCEU	P45805 alcaligenes
41	7	2.6	396	1	APA4_HUMAN	P06727 homo sapien
42	7	2.6	401	1	APA4_PAPAN	Q28758 papio anubi
43	7	2.6	412	1	AATC_BOVIN	P33097 bos taurus
44	7	2.6	412	1	AATC_BOVIN	P08906 equus cabal
45	7	2.6	412	1	AATC_HUMAN	P17174 homo sapien
46	7	2.6	412	1	AATC_MOUSE	P05201 mus musculu
47	7	2.6	412	1	AATC_PIG	P00503 sus scrofa
48	7	2.6	412	1	AATC_PIG	P13221 rattus norv
49	7	2.6	429	1	APA4_MACFA	P33621 macaca fasc
50	7	2.6	443	1	YJfJ_ECOLI	P39410 escherichia
51	7	2.6	473	1	PCC6_MYCTU	Q10506 mycobacteri
52	7	2.6	476	1	PCCB_RHOER	Q06101 rhodococcus
53	7	2.6	486	1	MVL_DROME	P49283 drosophila
54	7	2.6	504	1	FTSY_SYNY3	P73930 synechocyst
55	7	2.6	530	1	UDB2_RAT	P08541 rattus norv
56	7	2.6	534	1	FM2_ACTNA	P12616 actinomyces
57	7	2.6	538	1	NRM1_PIG	O77741 sus scrofa
58	7	2.6	548	1	AAK1_RAT	P54645 rattus norv
59	7	2.6	548	1	NRM1_BISBI	O95102 bison bison
60	7	2.6	548	1	NRM1_BOVIN	Q27981 bos taurus
61	7	2.6	548	1	NRM1_BUBBU	Q27946 bubalus bub
62	7	2.6	548	1	NRM1_CEREL	P56436 cervus elap
63	7	2.6	548	1	NRM1_MOUSE	P41251 mus musculu
64	7	2.6	548	1	NRM1_SHEEP	P49280 ovis aries
65	7	2.6	550	1	NRM1_HUMAN	P49279 homo sapien
66	7	2.6	568	1	NRM2_HUMAN	P49281 homo sapien
67	7	2.6	568	1	NRM2_MOUSE	P49282 mus musculu
68	7	2.6	568	1	NRM2_RAT	O54902 rattus norv
69	7	2.6	620	1	GG95_HUMAN	Q08379 homo sapien
70	7	2.6	629	1	VE1_HPVS1	P17382 human papil
71	7	2.6	653	1	CCME_RHOCA	Q00500 rhodobacter
72	7	2.6	726	1	PL01_BOVIN	O77588 bos taurus
73	7	2.6	790	1	SYFB_CHLMU	Q9pjf8 chlamydia m
74	7	2.6	834	1	TFIB_MOUSE	Q62318 mus musculu
75	7	2.6	835	1	TFIB_HUMAN	Q13263 homo sapien

## ALIGNMENTS

RESULT 1  
ID APAL\_HUMAN STANDARD: PRT: 267 AA.  
AC P02647;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=84221405; PubMed=6328445;  
RA Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,  
Baralle F.E.;  
RT "Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA  
abundance.";  
RL Nucleic Acids Res. 12:3917-3932(1984).  
[2]  
SEQUENCE FROM N.A.  
MEDLINE=85026665; PubMed=6207999;  
RA Selkhamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;  
RT "Isolation and DNA sequence of full-length cDNA and of the entire

RT gene for human apolipoprotein AI -- discovery of a new genetic  
 RL polymorphism in the apo AI gene.";  
 RN DNA 3:309-317(1984).  
 [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83220822; PubMed=6406984;  
 RA Shoulters C.C., Kornblitt A.R., Munro B.S., Baralle F.E.;  
 RT "Gene structure of human apolipoprotein AI.";  
 RL Nucleic Acids Res. 11:2827-2837(1983).  
 [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83220772; PubMed=6304641;  
 RA Cheung P., Chan L.;  
 RT "Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";  
 RL Nucleic Acids Res. 11:3703-3715(1983).  
 [15]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84119464; PubMed=6198645;  
 RA Law S.W., Brewer H.B. Jr.;  
 RT "Nucleotide sequence and the encoded amino acids of human  
 apolipoprotein A-I mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).  
 [16]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86008382; PubMed=2995392;  
 RA Law S.W., Brewer H.B. Jr.;  
 RT "Tangier disease. The complete mRNA sequence encoding for  
 preproapo-A-I.";  
 RL J. Biol. Chem. 260:12810-12814(1985).  
 [17]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84016011; PubMed=6413973;  
 RA Karathanasis S.K., Zannis V.I., Breslow J.L.;  
 RT "Isolation and characterization of the human apolipoprotein A-I  
 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).  
 [18]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89377481; PubMed=2673706;  
 RA Mogilevsky N., Roobol C., Lortiau R., Guillaume J.P., Jacobs P.,  
 RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,  
 RA Holmquist L., Carlson L.A., Bollen A.;  
 RT "Production of human recombinant proapolipoprotein A-I in Escherichia  
 coli: purification and biochemical characterization.";  
 RL DNA 8:429-436(1989).  
 [19]  
 RP SEQUENCE FROM N.A. (VARIANT TANGIER).  
 RX MEDLINE=88196137; PubMed=3129297;  
 RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,  
 RA Zannis V.I.;  
 RT "Sequence and expression of Tangier apoA-I gene.";  
 RL Eur. J. Biochem. 173:465-471(1988).  
 [110]  
 RP SEQUENCE OF 118-267 FROM N.A.  
 RX MEDLINE=83091059; PubMed=6294659;  
 RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,  
 RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;  
 RT "Isolation and characterization of cDNA clones for human  
 apolipoprotein A-I.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).  
 [111]  
 RP SEQUENCE OF 19-27.  
 RX MEDLINE=83256553; PubMed=6409108;  
 RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,  
 RA Light J.A.;  
 RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";  
 RL Biochem. Biophys. Res. Commun. 113:626-632(1983).  
 [112]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=78123731; PubMed=204308;  
 RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,  
 RA Bronzert T.J.;  
 RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated

RT from high density lipoproteins.";  
 RL Biochem. Biophys. Res. Commun. 80:623-630(1978).  
 [113]  
 RP SEQUENCE OF 25-267.  
 RX MEDLINE=75133493; PubMed=164450;  
 RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;  
 RT "The primary structure of human plasma high density apolipoprotein  
 glutamine I (APOA-I). II. The amino acid sequence and alignment of  
 RT cyanogen bromide fragments IV, III, and I.";  
 RL J. Biol. Chem. 250:2725-2738(1975).  
 [114]  
 RP SEQUENCE OF 25-56.  
 RX MEDLINE=88331387; PubMed=3047170;  
 RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;  
 RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein  
 A-I (Apo A-I). A novel function of Apo A-I.";  
 RL J. Clin. Invest. 82:803-807(1988).  
 [115]  
 RP SEQUENCE OF 25-48.  
 RX MEDLINE=89380318; PubMed=2506184;  
 RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,  
 RA Chapdelaine A.;  
 RT "Apolipoprotein A-I binds to a family of bovine seminal plasma  
 RT proteins.";  
 RL J. Biol. Chem. 264:16853-16857(1989).  
 [116]  
 RP SEQUENCE OF 25-43.  
 RX MEDLINE=88070603; PubMed=3120314;  
 RA Prioli R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,  
 RA Pereira M.E.A.;  
 RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi  
 RT neuraminidase, to high-density lipoprotein.";  
 RL Science 238:1417-1419(1987).  
 [117]  
 RP SEQUENCE OF 25-42.  
 RC TISSUE-Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 [118]  
 RP PALMITOYLATION.  
 RX MEDLINE=86140194; PubMed=3005308;  
 RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;  
 RT "Human apolipoprotein A-I. Post-translational modification by fatty  
 RT acid acylation.";  
 RL J. Biol. Chem. 261:3911-3914(1986).  
 [119]  
 RP PROCESSING.  
 RX MEDLINE=83195100; PubMed=6405383;  
 RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,  
 RA Breslow J.L.;  
 RT "Intracellular and extracellular processing of human apolipoprotein  
 RT A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).  
 [120]  
 RP STRUCTURE BY NMR OF 190-209.  
 RX MEDLINE=96270776; PubMed=8664326;  
 RA Wang G., Treleaven W.D., Cushley R.J.;  
 RT "Conformation of human serum apolipoprotein A-I(166-185) in the  
 RT presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR  
 and CD. Evidence for specific peptide-SDS interactions.";  
 RL Biochim. Biophys. Acta 1301:174-184(1996).  
 [121]  
 RP X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.  
 RX MEDLINE=98024124; PubMed=9356442;  
 RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;  
 RT "Crystal structure of truncated human apolipoprotein A-I suggests a  
 RT lipid-bound conformation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).  
 [122]  
 RP VARIANT MILANO.



FT REPEAT 123 144 4.  
FT REPEAT 145 166 5.  
FT REPEAT 167 188 6.  
FT REPEAT 189 210 7.  
FT REPEAT 211 232 8.  
FT REPEAT 233 243 9 (HALF-LENGTH).  
FT REPEAT 244 267 10.  
FT CONFLICT 13 L -> P (IN REF. 1).  
SQ SEQUENCE 267 AA; 30735 MW; 869955C024088E21 CRC64;

Query Match 25.8%; Score 69; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.5e-59;  
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 SKLRQLGPVTQEFWMDNLEKTEGLRQMSKDLKLVKAKVQPYLDDFQKKQWEMELLYRQ 141  
DB 82 SKLRQLGPVTQEFWMDNLEKTEGLRQMSKDLKLVKAKVQPYLDDFQKKQWEMELLYRQ 141  
OY 142 KVEPLRAEL 150  
DB 142 KVEPLRAEL 150

RESULT 3  
APAL\_TUPGB STANDARD; PRT; 265 AA.  
AC 018759;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Tupaia glis belangeri (Common tree shrew).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.  
OX NCBI\_TaxID=9396;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Lu X., Chen B., Zhao Y., Wang K., Xue H., Zeng W.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF005638; AAB82326.1; -.  
DR HSSP: P02647; IAV1.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam: PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18 BY SIMILARITY.  
FT PROPEP 19 24 BY SIMILARITY.  
FT CHAIN 25 265 APOLIPOPROTEIN A-I.  
FT DOMAIN 67 265 10 X APPROXIMATE TANDEM REPEATS.  
FT REPEAT 67 88 1.  
FT REPEAT 89 110 2.  
FT REPEAT 111 121 3 (HALF-LENGTH).  
FT REPEAT 122 143 4.  
FT REPEAT 144 165 5.

FT REPEAT 166 187 6.  
FT REPEAT 188 209 7.  
FT REPEAT 210 231 8.  
FT REPEAT 232 242 9 (HALF-LENGTH).  
FT REPEAT 243 265 10.  
SQ SEQUENCE 265 AA; 30332 MW; 60076BC39FAEA165 CRC64;

Query Match 8.6%; Score 23; DB 1; Length 265;  
Best Local Similarity 100.0%; Pred. No. 1.2e-14;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLTAVLFLTGSQARHFWQDEP 27  
DB 5 VLTAVLFLTGSQARHFWQDEP 27

RESULT 4  
APAL\_CANFA STANDARD; PRT; 266 AA.  
AC P02648;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=90132271; PubMed=2515239;  
RA Luo C.-C., Li W.-H., Chan L.;  
RT "Structure and expression of dog apolipoprotein A-I, E, and C-I  
RT mRNAs: implications for the evolution and functional constraints of  
RT apolipoprotein structure."  
RL J. Lipid Res. 30:1735-1746(1989).  
RN [2]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=82142425; PubMed=6801039;  
RA Chung H., Randolph A., Reardon I., Heinrichson R.L.;  
RT "The covalent structure of apolipoprotein A-I from canine high  
RT density lipoproteins."  
RL J. Biol. Chem. 257:2961-2967(1982).  
RN [3]  
RP SEQUENCE OF 25-57 AND 262-265.  
RX MEDLINE=76210910; PubMed=179887;  
RA Nakai T., Whayne T.F., Tang J.;  
RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein  
RT A-I."  
RL FEBS Lett. 64:409-411(1976).  
RN [4]  
RP SEQUENCE OF 25-37.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
RT dog heart proteins."  
RL Electrophoresis 18:2795-2802(1997).  
CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS. SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.  
CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
DR PIR: A03092; LPDGA1.  
DR HSSP: P02647; IODR.  
DR HSC-2DPAGE: P02648; DOG.  
DR InterPro: IPR000074; Apolipoprotein.



DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18 BY SIMILARITY.  
FT PROPEP 19 24  
FT CHAIN 25 266  
FT DOMAIN 67 266  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 231  
FT REPEAT 232 242  
FT REPEAT 243 266  
FT CONFLICT 168 168  
FT CONFLICT 202 202  
FT CONFLICT 235 235  
FT CONFLICT 264 266  
SQ SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;

Query Match 8.2%; Score 22; DB 1; Length 266;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LTLAVLFTGSGARHFWQDDP 27  
Db 6 LTLAVLFTGSGARHFWQDDP 27

RESULT 5  
APAL\_RABBIT STANDARD; PRT; 266 AA.  
AC P09809;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2ZAP AI; TISSUE=Small intestine;  
RA Paraskevopoulou T.B., Kritis A., Zannis V.I.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=88082866; PubMed=3121329;  
RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,  
RA Kroon P.A., Chao Y.S.;  
RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit  
RT apolipoprotein A-I is synthesized in the intestine but not in the  
RT liver.";  
RL Eur. J. Biochem. 170:99-104(1987).  
RN [3]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=87030294; PubMed=3095115;  
RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;  
RT "The primary structure of apolipoprotein A-I from rabbit high-density  
RT lipoprotein.";  
RL Eur. J. Biochem. 160:427-431(1986).

CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
CC THE LECTININ CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN

CC CHYLOMICRONS.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X15908; CAA34024.1; -  
DR EMBL; X06658; CAA29857.1; -  
DR EMBL; X06659; CAA29858.1; -  
DR PIR; S06064; LPRB1Z.  
DR PIR; S00230; LPRB1B.  
DR PIR; A24998; A24998.  
DR HSSP; P02647; LAV1.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 266  
FT DOMAIN 67 266  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 231  
FT REPEAT 232 242  
FT REPEAT 243 266  
FT CONFLICT 168 168  
FT CONFLICT 202 202  
FT CONFLICT 235 235  
FT CONFLICT 255 256  
SQ SEQUENCE 266 AA; 30591 MW; 0FF6DB386497C7D2 CRC64;

Query Match 8.2%; Score 22; DB 1; Length 266;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 86 EQLGPVTOEFWDNLEKTEGLR 107  
Db 85 EQLGPVTOEFWDNLEKTEGLR 106  
RESULT 6  
APAL\_BOVIN STANDARD; PRT; 265 AA.  
AC P15497;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90348478; PubMed=2117227;  
 RA O'Huigin C., Chan L., Li W.H.;  
 RT "Cloning and sequencing of bovine apolipoprotein A-I cDNA and  
 RL molecular evolution of apolipoproteins A-I and B-100.";  
 RN Mol. Biol. Evol. 7:327-339(1990).  
 RN (2)  
 RP SEQUENCE OF 19-265.  
 RX MEDLINE=92153895; PubMed=1739745;  
 RA Sparrow D.A., Lee B.R., Laplaud M.P., Auboliron S., Bauchart D.,  
 RA Chapman J.M., Gatto A.M. Jr., Yang C.Y., Sparrow J.T.;  
 RT "Plasma lipid transport in the pre-ruminant calf, Bos spp: primary  
 RL structure of bovine apolipoprotein A-I.";  
 RN Biochim. Biophys. Acta 1123:145-150(1992).  
 RN (3)  
 RP SEQUENCE OF 25-70.  
 RX MEDLINE=90147795; PubMed=2105728;  
 RA Auboliron S., Sparrow D.A., Beaubatie L., Bauchart D., Sparrow J.T.,  
 RA Laplaud M.P., Chapman J.M.;  
 RT "Characterization and amino-terminal sequence of apolipoprotein A-I  
 RL from plasma high density lipoproteins in the pre-ruminant calf, Bos  
 RT spp.";  
 RN Biochem. Biophys. Res. Commun. 166:833-839(1990).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL EFFLUX FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M35870; AAA30381.1; -.  
 DR PIR; A34649; A34649.  
 DR HSSP; P02647; 10DR.  
 DR InterPro: IPR000074; Apolipoprotein.  
 DR Pfam: PF01442; Apolipoprotein; 1.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 265  
 FT DOMAIN 67 265  
 FT REPEAT 67 88  
 FT REPEAT 89 110  
 FT REPEAT 111 121  
 FT REPEAT 122 143  
 FT REPEAT 144 165  
 FT REPEAT 166 187  
 FT REPEAT 188 209  
 FT REPEAT 210 230  
 FT REPEAT 231 241  
 FT REPEAT 242 265  
 FT CONFLICT 185 186  
 FT SEQUENCE 265 AA; 30276 MW; 06A2681EA2ABA50F CRC64;  
 Query Match 7.98; Score 21; DB 1; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 1e-12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VLTLAVLFLTGSQARHFWOOD 25  
 DB 5 VLTLAVLFLTGSQARHFWOOD 25  
 RESULT 7  
 ID APAL\_PIG STANDARD; PRT; 265 AA.  
 AC P18648;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OC NCBI\_TaxID=9823;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93224154; PubMed=8468059;  
 RA Birchbauer A., Knipping G., Juritsch B., Aschauer H., Zechner R.;  
 RT "Characterization of the apolipoprotein AI and CIII genes in the  
 RL domestic pig.";  
 RN Genomics 15:643-652(1993).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver, and Brain;  
 RX MEDLINE=94125128; PubMed=8294940;  
 RA Moeckel B., Zinke H., Flach R., Weiss B., Weller-Guettler H.,  
 RA Gassen H.;  
 RT "Expression of apolipoprotein A-I in porcine brain endothelium in  
 RL vitro.";  
 RN J. Neurochem. 62:788-798(1994).  
 RN (3)  
 RP SEQUENCE OF 34-265 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=90132667; PubMed=2105375;  
 RA Weller-Guettler H., Sommerfeldt M., Papandrikopoulou A., Mischek U.,  
 RA Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;  
 RT "Synthesis of apolipoprotein A-I in pig brain microvascular  
 RL endothelial cells.";  
 RN J. Neurochem. 54:444-450(1990).  
 RN (4)  
 RP SEQUENCE OF 105-265 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93154581; PubMed=8428656;  
 RA Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;  
 RT "Sequences and expression of the porcine apolipoprotein A-I and C-III  
 RN mRNAs.";  
 RN Gene 123:173-179(1993).  
 RN (5)  
 RP SEQUENCE OF 25-265.  
 RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,  
 RA Jackson K., Gustavsson I., Rapacz J.;  
 RL Submitted (OCT-1995) to the SWISS-PROT data bank.  
 RN (6)  
 RP SEQUENCE OF 25-34.  
 RX MEDLINE=76184721; PubMed=178359;  
 RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;  
 RT "Characterization of the plasma lipoproteins and apoproteins of the  
 RL Erythrocebus patas monkey.";  
 RN Biochemistry 15:1928-1933(1976).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE  
 CC LIVER.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L00626; AAA30992.1; -;  
DR EMBL; X69477; CAA49234.1; -;  
DR EMBL; X17057; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X59414; CAA42050.1; -;  
DR PIR; JQ0704; JQ0704.  
DR PIR; A05311; A05311.  
DR PIR; A46018; A46018.  
DR PIR; S21830; S21830.  
DR PIR; S31394; S31394.  
DR HSSP; P02647; 10DR.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL; 1 18  
FT PROPEP 19 24  
FT CHAIN 25 265  
FT DOMAIN 67 265  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 142  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 230  
FT REPEAT 231 241  
FT REPEAT 242 265  
FT CONFLICT 108 108  
FT CONFLICT 143 143  
FT CONFLICT 173 173  
FT CONFLICT 180 180  
FT CONFLICT 185 186  
FT CONFLICT 209 209  
FT CONFLICT 224 224  
SQ SEQUENCE 265 AA; 30325 MW; 2C6E578318ECF69C CRC64;

Query Match 7.9%; Score 21; DB 1; Length 265;  
Best Local Similarity 100.0%; Pred. No. 1e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLTIAVLFLTGSQARHFWOOD 25  
|||  
Db 5 VLTIAVLFLTGSQARHFWOOD 25

RESULT 8  
APAI\_RAT  
ID APAI\_RAT STANDARD; PRT; 259 AA.  
AC P04639;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN APOA1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84207987; PubMed=6426956;  
RA Poncin J.E., Martial J.A., Gielen J.E.;  
RT "Cloning and structure analysis of the rat apolipoprotein A-I cDNA.";  
RL Eur. J. Biochem. 140:493-498(1984).

RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87008540; PubMed=3020028;  
RA Haddad I.A., Ordovas J.M., Fitzpatrick T., Karathanasis S.K.;  
RT "Linkage, evolution, and expression of the rat apolipoprotein A-I, C-  
RT III, and A-IV genes.";  
RL J. Biol. Chem. 261:13268-13277(1986).

RN [3]  
RP SEQUENCE OF 1-45.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=82098162; PubMed=6798036;  
RA Gordon J.I., Smith D.P., Andy R., Alpers D.H., Schonfeld G.,  
RA Strauss A.W.;  
RT "The primary translation product of rat intestinal apolipoprotein A-I  
RT mRNA is an unusual preprotein.";  
RL J. Biol. Chem. 257:971-978(1982).  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M00001; AAA40749.1; -;  
DR EMBL; X00558; CAA25224.1; -;  
DR EMBL; J02597; AAA40745.1; -;  
DR PIR; A24700; A24700.  
DR HSSP; P02647; 1A-V1.  
DR InterPro: IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL; 1 18  
FT PROPEP 19 24  
FT CHAIN 25 259  
FT DOMAIN 67 259  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 161  
FT REPEAT 162 183  
FT REPEAT 184 203  
FT REPEAT 204 225  
FT REPEAT 226 236  
FT REPEAT 237 259  
FT CONFLICT 201 201  
FT CONFLICT 214 214  
FT CONFLICT 218 218  
SQ SEQUENCE 259 AA; 30088 MW; 2E8D5EB45FEAE88 CRC64;

Query Match 4.5%; Score 12; DB 1; Length 259;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KDSGRDYVSQFE 58  
|||  
Db 46 KDSGRDYVSQFE 57

RESULT 9  
APAI\_MOUSE  
ID APAI\_MOUSE STANDARD; PRT; 264 AA.

AC Q00623; 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92281682; PubMed=1596360;  
 RA Stoffel W., Mueller R., Binczek E., Hofmann K.;  
 RT "Mouse apolipoprotein AI. cDNA-derived primary structure, gene  
 RT organisation and complete nucleotide sequence.";  
 RL Biol. Chem. Hoppe-Seyler 373:187-193(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93122774; PubMed=1478650;  
 RA Janusz J.L., Azrolan N., O'Connell A., Aalto-Setälä K., Breslow J.L.;  
 RT "Characterization of the mouse apolipoprotein ApoA-1/Apoc-3 gene  
 RT locus: genomic, mRNA, and protein sequences with comparisons to other  
 RT species.";  
 RL Genomics 14:1081-1088(1992).  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X64262; CAA45560.1; -;  
 DR EMBL; X64263; CAA45561.1; -;  
 DR EMBL; L04149; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; L04151; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S22420; S22420.  
 DR PIR; A44364; A44364.  
 DR HSSP; P02647; 1AV1.  
 DR SWISS-2DPAGE; Q00623; MOUSE.  
 DR MGD; MGI:88049; APOA1.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 1.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24 BY SIMILARITY.  
 FT CHAIN 25 264 BY SIMILARITY.  
 FT DOMAIN 67 264 APOLOPOPROTEIN A-I.  
 FT REPEAT 67 88 10 X APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 89 110 1.  
 FT REPEAT 111 121 2.  
 FT REPEAT 122 143 3 (HALF-LENGTH).  
 FT REPEAT 144 165 4.  
 FT REPEAT 166 187 5.  
 FT REPEAT 188 207 6.  
 FT REPEAT 208 229 7 (INCOMPLETE).  
 FT REPEAT 230 240 8.  
 FT REPEAT 241 264 9 (HALF-LENGTH).  
 SQ SEQUENCE 264 AA; 30587 MW; C453FF2019634AAC CRC64;

Query Match 4.5%; Score 12; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 0.00054;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 47 KDSGRVYSQFE 58  
 Db 46 KDSGRVYSQFE 57  
 RESULT 10  
 ID APOA1-ANAPL STANDARD; PRT; 264 AA.  
 AC 042296;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Apolipoprotein A-I precursor (Apo-AI).  
 GN APOA1.  
 OS Anas platyrhynchos (Domestic duck).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.  
 OC NCBI\_TaxID=8839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEKIN BREED; TISSUE=Liver;  
 RA Chen B., Lu X., Wang K., Xue H., Zeng W., Zhu D., Chi L.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR  
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
 CC CHYLOMICRONS.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U86131; AAB64381.1; -;  
 DR HSSP; P02647; 1ODR.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR Pfam; PF01442; Apolipoprotein; 1.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24 BY SIMILARITY.  
 FT CHAIN 25 264 APOLOPOPROTEIN A-I.  
 FT DOMAIN 67 264 10 X APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 67 88 1.  
 FT REPEAT 89 110 2.  
 FT REPEAT 111 121 3 (HALF-LENGTH).  
 FT REPEAT 122 143 4.  
 FT REPEAT 144 165 5.  
 FT REPEAT 166 187 6.  
 FT REPEAT 188 209 7.  
 FT REPEAT 210 231 8.  
 FT REPEAT 232 242 9 (HALF-LENGTH).  
 FT REPEAT 243 264 10.  
 SQ SEQUENCE 264 AA; 30659 MW; 7546F4BAF2846900 CRC64;

Query Match 3.7%; Score 10; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 LAPYSDELRO 196  
 Db 186 LAPYSDELRO 195



```

RESULT 11
APAL_CHICK STANDARD; PRT; 264 AA.
ID APAL_CHICK
AC P08250;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88049703; PubMed=3118875;
RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.;
RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and
RL evolution."
RN [2]
RP Biochem. Biophys. Res. Commun. 148:485-492(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88152500; PubMed=3126099;
RA Ferrari S., Tarugi P., Drusiani E., Calandra S., Fregni M.;
RT "The complete sequence of chick apolipoprotein AI mRNA and its
RN expression in the developing chick."
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=87222301; PubMed=3108248;
RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,
RA Leberer H., Lusis A.J.;
RT "Structure, evolution, and regulation of chicken apolipoprotein A-I."
RN J. Biol. Chem. 262:7058-7065(1987).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92381402; PubMed=1512510;
RA Lamou-Fava S., Sasstry R., Ferrari S., Rajavashisth T.B.,
RA Lusis A.J., Karathanasis S.K.;
RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene
RN expression: differences between avian and mammalian apoA-I gene
RL transcription control regions."
RN J. Lipid Res. 33:831-842(1992).
RN [6]
RP SEQUENCE OF 25-44.
RX MEDLINE=83213468; PubMed=6406496;
RA Shackelford J.E., Leberer H.G.;
RT "Synthesis and secretion of apolipoprotein A1 by chick breast
RN muscle."
RN J. Biol. Chem. 258:7175-7180(1983).
RN [7]
RP FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
RX CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
RN THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
RN [8]
RP SUBCELLULAR LOCATION: Extracellular.
RN [9]
RP TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
RN CHYLOMICRONS.
RN [10]
RP SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
RN [11]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17961; AAA48593.1; -
DR EMBL; M18746; AAA48594.1; -
DR EMBL; M25559; AAA48592.1; -
DR EMBL; M96012; AAA48597.1; -
DR PIR; S01453; LPCHA1.

```

```

DR PIR; JH0471; JH0471.
DR HSSP; P02647; 1ODR.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein. 1.
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 264
FT DOMAIN 67 264
FT REPEAT 67 88
FT REPEAT 89 110
FT REPEAT 111 121
FT REPEAT 122 143
FT REPEAT 144 165
FT REPEAT 166 187
FT REPEAT 188 209
FT REPEAT 210 231
FT REPEAT 232 242
FT REPEAT 243 264
FT CONFLICT 16 16
FT CONFLICT 148 148
FT SEQUENCE 264 AA; 30680 MW; 968320E81E2AC5C2 CRC64;

```

```

Query Match 3.7%; Score 10; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 LADYSEDLRQ 196
Db 186 LADYSEDLRQ 195

```

```

RESULT 12
APAL_COTJA STANDARD; PRT; 264 AA.
ID APAL_COTJA
AC P32918;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN APOA1.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RN MEDLINE=97212027; PubMed=9058967;
RA Oku H., Toda T., Nagata J., Ishikawa M., Neyazaki K., Shinjo C.,
RA Chinen I.;
RT "Apolipoprotein A-1 of Japanese quail: cDNA sequence and modulation of
RN tissue expression by cholesterol feeding."
RN Biosci. Biotechnol. Biochem. 61:286-290(1997).
RN [2]
RP SEQUENCE OF 25-60.
RX MEDLINE=93213845; PubMed=8461329;
RA Oku H., Ishikawa M., Nagata J., Toda T., Chinen I.;
RT "Lipoprotein and apoprotein profile of Japanese quail."
RN Biochim. Biophys. Acta 1167:22-28(1993).
RN [3]
RP FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
RX CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
RN THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
RN [4]
RP SUBCELLULAR LOCATION: Extracellular.
RN [5]
RP TISSUE SPECIFICITY: MAJOR PROTEIN OF VLDL, HDL, LDL AND IN
RN CHYLOMICRONS. EXPRESSED IN A NUMBER OF TISSUES INCLUDING LIVER,
CC SMALL INTESTINE, LUNG, KIDNEY, HEART AND MUSCLE WITH HIGHEST
CC EXPRESSION IN LIVER AND SMALL INTESTINE.
CC [6]
RP SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; D85133; BAA12729.1; -.  
DR HSSP; P02647; 1GWA.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 264  
FT DOMAIN 67 264  
FT REPEAT 67 88  
FT REPEAT 89 110  
FT REPEAT 111 121  
FT REPEAT 122 143  
FT REPEAT 144 165  
FT REPEAT 166 187  
FT REPEAT 188 209  
FT REPEAT 210 231  
FT REPEAT 232 242  
FT REPEAT 243 264  
SQ SEQUENCE 264 AA; 30753 MW; 8781DE213C3F863F CRC64;

Query Match 3.4%; Score 9; DB 1; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TLAVLFTG 15  
| | | | |  
Db 7 TLAVLFTG 15

RESULT 13  
APAL\_ERYPA STANDARD; PRT; 20 AA.  
ID APAL\_ERYPA  
AC P18647;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Apolipoprotein A-I (Apo-AI) (Fragment).  
GN APOA1.  
OS Erythrocebus patas (Red guenon) (Hussar).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Erythrocebus.  
OX NCBI\_TaxID=9538;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=76184721; PubMed=178359;  
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;  
RT "Characterization of the plasma lipoproteins and apoproteins of the  
RT Erythrocebus patas monkey.";  
RL Biochemistry 15:1928-1933(1976).  
CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF  
CC CHOLESTEROL EFFLUX FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING  
CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN  
CC CHYLOMICRONS.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
DR PIR; A05313; A05313.  
KW Plasma; Lipid transport; HDL; Cholesterol metabolism.  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2387 MW; 9C970997C7FC976A CRC64;

Query Match 3.0%; Score 8; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 PWDVKDL 38  
| | | | |  
Db 7 PWDVKDL 14

RESULT 14  
APE\_PIG STANDARD; PRT; 317 AA.  
ID APE\_PIG  
AC P18650; O19099;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Apolipoprotein E precursor (Apo-E).  
GN APOE.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Brzowska A.M., Grimholt U., Kulseth M.A., Wold I., Rognes S.;  
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98347372; PubMed=9682450;  
RA Ramsoondar J.J., Rucker E.B., Vasquez J.C., Gallagher D.S.,  
RA Grimm D.R., Lunney J.K., Schook L.B., Piedrahita J.A.;  
RT "Isolation and genetic characterization of the porcine apolipoprotein  
RT E gene.";  
RT Anim. Genet. 29:43-47(1998).  
RN [3]  
RP PRELIMINARY SEQUENCE OF 19-31.  
RX MEDLINE=81021043; PubMed=7417263;  
RA Weisgraber K.H., Troxler R.F., Rall S.C., Mahley R.W.;  
RT "Comparison of the human, canine and swine E apoproteins.";  
RL Biochem. Biophys. Res. Commun. 95:374-380(1980).  
CC -1- FUNCTION: APO-E MEDIATES BINDING, INTERNALIZATION, AND CATABOLISM  
CC OF LIPOPROTEIN PARTICLES. IT CAN SERVE AS A LIGAND FOR THE LDL(APO  
CC B/E) RECEPTOR AND FOR THE SPECIFIC APO-E RECEPTOR (CHYLOMICRON  
CC REMNANT) OF HEPATIC TISSUES.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.  
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; X72835; CAA51356.1; -.  
DR EMBL; U70240; AAC29512.1; -.  
DR PIR; A05312; A05312.  
DR PIR; S33450; S33450.  
DR HSSP; P02649; 1OEF.  
DR InterPro; IPR000074; Apolipoprotein.  
DR Pfam; PF01442; Apolipoprotein; 1.  
KW Glycoprotein; Plasma; Lipid transport; HDL; VLDL; Chylomicron;  
KW Heparin-binding; Repeat; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 317  
FT DOMAIN 157 167  
FT DOMAIN 161 164  
FT DOMAIN 228 235  
FT DOMAIN 254  
SQ SEQUENCE 254 AA; 254 MW; 8 X 22 AA APPROXIMATE TANDEM REPEATS.

```
FT REPEAT 79 100 1.
FT REPEAT 101 122 2.
FT REPEAT 123 144 3.
FT REPEAT 145 166 4.
FT REPEAT 167 188 5.
FT REPEAT 189 210 6.
FT REPEAT 211 232 7.
FT REPEAT 233 254 8.
FT CONFLICT 35 35 S -> P (IN REF. 2).
FT CONFLICT 160 161 NV -> KL (IN REF. 2).
FT CONFLICT 166 166 V -> L (IN REF. 2).
FT CONFLICT 251 252 DE -> EQ (IN REF. 2).
FT CONFLICT 273 273 A -> G (IN REF. 2).
FT CONFLICT 275 275 Q -> H (IN REF. 2).
FT CONFLICT 277 277 R -> L (IN REF. 2).
FT CONFLICT 289 289 M -> I (IN REF. 2).
FT CONFLICT 303 303 A -> G (IN REF. 2).
SQ SEQUENCE 317 AA; 36599 MW; 83E7F51A07785055 CRC64;
```

Query Match  
Best Local Similarity 100.0%; Score 8; DB 1; Length 317;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 QLGPTQ 94  
DB 98 QLGPTQ 105

```
RESULT 15
AP04_MOUSE STANDARD; PRT; 395 AA.
ID AP04_MOUSE
AC P06728;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV).
GN APOA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089722; PubMed=3796595;
RA Williams S.C., Bruckheimer S.M., Lusis A.J., LeBoeuf R.C.,
RA Kiniburgh A.J.;
RT "Mouse apolipoprotein A-IV gene: nucleotide sequence and induction by
RT a high-lipid diet."
RL Mol. Cell. Biol. 6:3807-3814(1986).
RN [2]
RP REVISIONS.
RA Kiniburgh A.J.;
RL Submitted (DEC-1986) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, AND 129/J;
RX MEDLINE=91286309; PubMed=1648102;
RA Reue K., Leete T.H.;
RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and
RT deletion in a region of tandem repeats."
RL J. Biol. Chem. 266:12715-12721(1991).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECTIN:CHOLESTEROL
```

```
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- POLYMORPHISM: THERE IS A POLYMORPHISM WITHIN A SERIES OF IMPERFECT
CC REPEATS ENCODING THE SEQUENCE E-Q-[AV]-Q. INSERTIONS OR DELETIONS
CC OF 12 NUCLEOTIDES HAVE GIVEN RISE TO THREE FORMS CHARACTERIZED BY
CC THREE (129), FOUR (C57BL/6), OR FIVE (M.CASTANEUS) COPIES OF THE
CC REPEAT UNIT.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13966; AAA37253.1; -
DR EMBL; M64249; AAA37215.1; -
DR EMBL; M64248; AAA37214.1; -
DR PIR; A25281; A25281.
DR MGI; 88051; APOA4.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;
KW Polymorphism.
FT SIGNAL 1 20
FT CHAIN 21 395
FT DOMAIN 33 330
FT REPEAT 33 54
FT REPEAT 60 81
FT REPEAT 82 103
FT REPEAT 115 136
FT REPEAT 137 158
FT REPEAT 159 180
FT REPEAT 181 202
FT REPEAT 203 224
FT REPEAT 225 246
FT REPEAT 247 268
FT REPEAT 269 286
FT REPEAT 287 308
FT REPEAT 309 330
FT DOMAIN 366 389
FT VARIANT 382 385
FT CONFLICT 15 15
FT CONFLICT 63 63
FT CONFLICT 207 207
FT CONFLICT 288 288
FT CONFLICT 294 295
FT CONFLICT 315 316
SQ SEQUENCE 395 AA; 45029 MW; 5FE27D0236226257 CRC64;
GLU/GLN-RICH.
MISSING (IN SOME STRAINS).
O -> K (IN REF. 1).
E -> R (IN REF. 1).
S -> A (IN REF. 1).
RQ -> KA (IN REF. 1).
NK -> GG (IN REF. 1).
```

Query Match  
Best Local Similarity 100.0%; Score 8; DB 1; Length 395;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAAVLTIA 9  
DB 4 KAAVLTIA 11

```
RESULT 16
UBR1_YEAST STANDARD; PRT; 1950 AA.
ID UBR1_YEAST
AC P19812;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE N-end-recognition protein (ubiquitin-protein ligase E3 component) (N-
DE recognin).
GN UBR1 OR PPR1 OR YGR184C OR G7168.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
```

```

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / GTF88;
RX MEDLINE=91006011; PubMed=2209542;
RA Bartel B., Wuening I., Varshavsky A.;
RT "The recognition component of the N-end rule pathway.";
RL EMO J. 9:3179-3189(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97279231; PubMed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
of Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:357-363(1997).
RN [3]
RP SEQUENCE OF 346-1950 FROM N.A.
RA Hebling U., Hofmann B., Delius H.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECOGNITION COMPONENT OF THE N-END RULE PATHWAY. BINDS
CC TO PROTEINS BEARING AMINO-TERMINAL RESIDUES THAT ARE DESTABILIZING
CC ACCORDING TO THE N-END RULE, BUT DOES NOT BIND TO OTHERWISE
CC IDENTICAL PROTEINS BEARING STABILIZING AMINO-TERMINAL RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: X53747; CAA37779.1; -.
CC EMBL: X99074; CAA67528.1; -.
CC EMBL: Z72969; CAA97210.1; -.
CC PIR: S12332; S12332.
CC SGD: S0003416; UBR1.
CC InterPro: IPR003126; znf_UBR1.
CC InterPro: IPR001841; znf_ring.
CC Pfam: PF02207; zf-UBR1; 1.
CC SMART: SM00184; RING; 1.
CC SMART: SM00396; znf_UBR1; 1.
CC Ligase; Ubiquitin conjugation.
CC KW SEQUENCE 1950 AA; 224836 MW; 33E4CD3A031AF523 CRC64;
SQ

```

```

RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -I- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002297; AAF73544.1; -.
DR TIGR; TC0295; -.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00886; Ribosomal_S16; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 116 AA; 13452 MW; 298132809F2C5011 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. NO. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 RORLAAR 201
DB 103 RORLAAR 109
|||||||

RESULT 18
RS16_CHLTR
ID RS16_CHLTR STANDARD; PRT; 116 AA.
AC 084029;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S16.
GN RPS16 OR RS16 OR CT026.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -I- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001277; AAC67616.1; -.
DR HSSP; P80379; 1EMW.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00886; Ribosomal_S16; 1.

```

DR PROSITE; PS00732; RIBOSOMAL\_S16; FALSE\_NEG.  
KW Ribosomal protein; complete proteome.  
SQ SEQUENCE 116 AA; 13409 MW; 298132991EB44721 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 QRLAAR 201  
Db 103 QRLAAR 109

RESULT 19

TRMA\_LACLA

ID TRMA\_LACLA STANDARD; PRT; 132 AA.

AC Q9ETA4;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Thermoresistant insertional mutant protein A.

GN TRMA OR NRPR OR L0675.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis), and

OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Lactococcus.

OX NCBI\_TaxID=1360, 1359;

RN [1]

SEQUENCE FROM N.A.

RC SPECIES=L.1.lactis; STRAIN=IL1403;

RX MEDLINE=21235186; PubMed=11337471;

RA Bolotin A., Winkler P., Manger S., Jallou O., Malarme K.,

RA Weissbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus

RT lactis ssp. lactis IL1403."

RL Genome Res. 11:731-753(2001).

RN [2]

SEQUENCE FROM N.A.

RC SPECIES=L.1.cremoris; STRAIN=MG1363;

RX MEDLINE=99157588; PubMed=10048028;

RA Duwat P., Ehrlich S.D., Gruss A.;

RT "Effects of metabolic flux on stress response pathways in Lactococcus

RT lactis."

RL Mol. Microbiol. 31:845-858(1999).

RN [3]

SEQUENCE FROM N.A.

RC SPECIES=L.1.cremoris; STRAIN=MG1363;

RA Frees D., Ingmer H.I.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE ARSC FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AE006301; AAK04773.1; -

DR EMBL; AF058950; AAG43111.1; -

DR EMBL; AF028804; AAG37358.1; -

KW Oxidoreductase; Complete proteome.

SQ SEQUENCE 132 AA; 15030 MW; F63EE4B8204136B3 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 ENGARL 213  
|||||||

Db 122 ENGARL 128

RESULT 20

YB02\_AERPE

ID YB02\_AERPE STANDARD; PRT; 139 AA.

AC Q9YD10;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein APE1102.

GN APE1102.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;

OC Aeropyrum.

OX NCBI\_TaxID=56636;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=KI;

RX MEDLINE=99310339; PubMed=10382966;

RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,

RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,

RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

RT "Complete genome sequence of an aerobic hyper-thermophilic

RT crenarchaeon, Aeropyrum pernix KI."

RL DNA Res. 6:83-101(1999).

CC -!- SIMILARITY: BELONGS TO THE UPF0129 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AP000060; BAA80087.1; -

DR InterPro; IPR002851; DUF133.

DR Pfam; PF02000; DUF133; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 139 AA; 14804 MW; 0F16DEEDDC88686 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 139;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAAR 202  
Db 120 QRLAAR 126

RESULT 21

YIGI\_ECOLI

ID YIGI\_ECOLI STANDARD; PRT; 155 AA.

AC P27845; P76761;

DT 01-AUG-1992 (Rel. 23, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein yigI.

GN YIGI OR B3820.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxID=562;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=92358234; PubMed=1379743;

RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;



RP "Analysis of the Escherichia coli genome: DNA sequence of the region  
RT from 84.5 to 86.5 minutes."  
RL Science 257:771-778(1992).  
RN [12]  
RP REVISION TO 96.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
RN [13]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RA Ohmori H.;  
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP SEQUENCE OF 1-42 FROM N.A.  
RX MEDLINE=85157492; PubMed=6397464;  
RA Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,  
RA Inoue K., Ikeda H., Sekiguchi M., Nojima S.;  
RT "The DNA sequence encoding pldA gene, the structural gene for  
RT detergent-resistant phospholipase A of E. coli.";  
RL J. Biochem. 96:1655-1664(1984).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M87049; AAA67616.1; ALT\_INIT.  
DR EMBL; AE000457; AAC76823.1; ALT\_INIT.  
DR EMBL; L02122; AAD15042.1; ALT\_INIT.  
DR EMBL; X02143; -; NOT\_ANNOTATED\_CDS.  
DR PIR; S30710; S30710.  
DR PIR; S30747; S30747.  
DR EcoGene; EG11467; YigI.  
DR InterPro; IPR003736; DUF157.  
DR Pfam; PF02584; DUF157; 1.  
KW Hypothetical protein; Complete proteome.  
FT CONFLICT 96 L -> V (IN REF. 1).  
SQ SEQUENCE 155 AA; 17163 MW; 96790B719B1E0609 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 155;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 192 DELRORL 198  
Db 90 DELRORL 96

RESULT 22  
YIGI\_SALTY STANDARD; PRT; 155 AA.  
ID YIGI\_SALTY  
AC P40725; Q9L6P0;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein yigI.  
GN YIGI OR STM3956 OR STWDI.34 OR STY3603.  
OS Salmonella typhimurium, and  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602, 601;  
RN [1]

RP SEQUENCE FROM N.A.  
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";  
RL Nature 413:852-856(2001).  
RN [12]  
RP SEQUENCE OF 1-25 FROM N.A.  
RC SPECIES=S.typhimurium;  
RX MEDLINE=94131966; PubMed=8300539;  
RA Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,  
RA Verheij H.M., Tommassen J.;  
RT "Molecular characterization of enterobacterial pldA genes encoding  
RT outer membrane phospholipase A.";  
RL J. Bacteriol. 176:861-870(1994).  
RN [13]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.typhi; STRAIN=CT18;  
RX MEDLINE=21534947; PubMed=11677608;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
RA Whitehead S., Barrett B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
RT enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF233324; AAF3436.1; -;  
DR EMBL; AE008884; AAL22800.1; ALT\_INIT.  
DR EMBL; X76900; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; AL627278; CAD07936.1; -;  
DR StyGene; SG10461; YigI.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 155 AA; 17043 MW; 546D5A3840701F2B CRC64;

Query Match 2.6%; Score 7; DB 1; Length 155;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 192 DELRORL 198  
Db 90 DELRORL 96

RESULT 23  
LOLB\_ACTAC STANDARD; PRT; 207 AA.  
ID LOLB\_ACTAC  
AC O52727;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Outer-membrane lipoprotein lolB precursor.  
GN LOLB OR HEMM.  
OS Actinobacillus actinomycetemcomitans (Haemophilus  
OS actinomycetemcomitans).



```

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUNY 465;
RA Mintz K.P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLAYS A CRITICAL ROLE IN THE INCORPORATION OF
CC LIPOPROTEINS IN THE OUTER MEMBRANE AFTER THEY ARE RELEASED BY
CC THE LOLA PROTEIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (By similarity).
CC -1- SIMILARITY: BELONGS TO THE LOLB FAMILY.
CC -1- CAUTION: WAS ORIGINALLY THOUGHT TO BE INVOLVED IN DELTA-
CC AMINOLEVULINIC ACID BIOSYNTHESIS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF045460; AAC04857.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Chaperone; Outer membrane; Lipoprotein; Transport; Protein transport;
KW Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 207 OUTER-MEMBRANE LIPOPROTEIN LOLB.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE 207 AA; 24219 MW; 5CE534F72F0CFB8D CRC64;

Query Match
Best Local Similarity 2.6%; Score 7; DB 1; Length 207;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 LAVLELT 14
Db 12 LAVLELT 18

RESULT 24
EUTL_ECOLI STANDARD; PRT; 219 AA.
AC P76541;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ethanolamine utilization protein eutL.
GN EUTL OR B2439.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: MAY BE INVOLVED IN THE FORMATION OF A SPECIFIC
CC MICROCOMPARTMENT IN THE CELL IN WHICH THE METABOLISM OF
CC POTENTIALLY TOXIC BY-PRODUCTS TAKES PLACE.
CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.
CC -1- SIMILARITY: DISTANTLY RELATED TO THE CCHA/CCMK/CSOS1/PDUA FAMILY.
-----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000331; AAC75492.1; -
DR Ecogene; EG14171; eutL.
KW Complete proteome.
SQ SEQUENCE 219 AA; 22788 MW; 1A5417565256243E CRC64;

Query Match
Best Local Similarity 2.6%; Score 7; DB 1; Length 219;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 FLTGSQA 18
Db 189 FLTGSQA 195

RESULT 25
EUTL_SALTY STANDARD; PRT; 219 AA.
AC Q92FU9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ethanolamine utilization protein eutL.
GN EUTL OR STM2456 OR STY2693.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RX MEDLINE=99395039; PubMed=10464203;
RA Kofoed E.C., Rappleye C.A., Stojiljkovic I., Roth J.R.;
RT "The 17-gene ethanolamine (eut) operon of Salmonella typhimurium
RL encodes five homologues of carboxysome shell proteins.";
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RL LT2.";
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RL enterica serovar Typhi CT18.";
RN [4]
RP "Nature 413:848-852(2001).";
CC -1- FUNCTION: MAY BE INVOLVED IN THE FORMATION OF A SPECIFIC
-----

```

CC MICROCOMPARTMENT IN THE CELL IN WHICH THE METABOLISM OF  
CC POTENTIALLY TOXIC BY-PRODUCTS TAKES PLACE.  
CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.  
CC -1- SIMILARITY: DISTANTLY RELATED TO THE CCHA/CCMK/CSOS1/PDUA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF093749; AAC78125.1; -  
CC EMBL; AE008810; AAL21350.1; -  
CC EMBL; AL627274; CAD07687.1; -  
CC StyGene; SG10631; eutl.  
CC Complete proteome.  
CC KW Complete proteome.  
CC SQ SEQUENCE 219 AA; 22695 MW; DA50E6EDF6FCD858 CRC64;  
  
Query Match 2.6%; Score 7; DB 1; Length 219;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 12 FLTGSQA 18  
|111111|  
Db 189 FLTGSQA 195  
  
RESULT 26  
YGI\_X\_ECOLI STANDARD; PRT; 219 AA.  
ID YGI\_X\_ECOLI  
AC P52076;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Probable transcriptional regulatory protein ygiX.  
GN YGI\_X OR B3025.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
CC -1- FUNCTION: PROBABLE MEMBER OF A TWO-COMPONENT REGULATORY SYSTEM  
CC YGI\_X/YGI\_Y.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- PTM: PHOSPHORYLATED BY YGI\_Y (POTENTIAL).  
CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE H1708.  
CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U28377; AAA60193.1; -  
CC EMBL; AE000384; AAC76061.1; -  
CC HSSP; P08402; 1B00.  
CC ECOSys; EG13026; YgiX.  
CC InterPro; IPR001789; Response\_reg.

DR InterPro; IPR001867; Trans\_reg\_C.  
DR Pfam; PF00072; response\_reg; 1.  
DR Pfam; PF00486; trans\_reg\_C; 1.  
DR SMART; SM00448; REC; 1.  
DR PROSITE; PS50110; RESPONSE REGULATORY; 1.  
KW Hypothetical protein; Sensory transduction; Phosphorylation;  
KW Transcription regulation; DNA-binding; Complete proteome.  
FT DOMAIN 1 116 RESPONSE REGULATORY.  
FT MOD\_RES 51 51 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 219 AA; 24677 MW; 9408B2B7261E042E CRC64;  
  
Query Match 2.6%; Score 7; DB 1; Length 219;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 199 AARLEAL 205  
|111111|  
Db 109 AARLEAL 115  
  
RESULT 27  
CHIT\_DIOJA STANDARD; PRT; 250 AA.  
ID CHIT\_DIOJA  
AC P80052;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Acidic endochitinase (EC 3.2.1.14).  
OS Dioscorea japonica (Yam).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;  
OC Dioscorea.  
OX NCBI\_TaxID=4673;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Aerial tuber;  
RX MEDLINE=93015853; PubMed=1400311;  
RA Araki T., Funatsu J., Kuramoto M., Konno H., Torikata T.;  
RT "The complete amino acid sequence of yam (Dioscorea japonica)  
RT chitinase. A newly identified acidic class I chitinase.";  
RL J. Biol. Chem. 267:19944-19947(1992).  
RN [2]  
RP SEQUENCE OF 1-53.  
RC TISSUE=Aerial tuber;  
RX MEDLINE=92322966; PubMed=1623187;  
RA Araki T., Funatsu J., Kuramoto M., Torikata T.;  
RT "Amino acid sequence of the N-terminal domain of yam (Dioscorea  
RT japonica) aerial tuber acidic chitinase. Evidence for the presence of  
RT a wheat germ agglutinin domain in matured acidic chitinase from  
RT unstressed tuber.";  
RL Plant Mol. Biol. 19:351-354(1992).  
CC -1- FUNCTION: DEFENSE AGAINST CHITIN CONTAINING FUNGAL PATHOGENS.  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
CC acetyl-D-glucosamine polymers of chitin.  
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS IA WHICH IS SIMILAR TO  
CC CLASS IB IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL  
CC CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL  
CC HYDROLASES).  
CC PIR; A40173; A40173.  
DR HSSP; P10969; 1WGT.  
DR InterPro; IPR001002; Chitin\_bind.  
DR InterPro; IPR000726; Glyco\_hydro\_19.  
DR Pfam; PF00187; chitin\_binding; 1.  
DR Pfam; PF00182; Glyco\_hydro\_19; 1.  
DR ProDom; PD000574; Glyco\_hydro\_19; 1.  
DR ProDom; PD000609; Chitin\_bind; 1.  
DR SMART; SM00270; ChtBD1; 1.  
DR PROSITE; PS00026; CHITIN\_BINDING; 1.  
DR PROSITE; PS00773; CHITINASE\_19\_1; 1.  
DR PROSITE; PS00774; CHITINASE\_19\_2; 1.  
KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT DOMAIN 1 36 CHITIN-BINDING (BY SIMILARITY).  
FT DISULFID 3 12 BY SIMILARITY.  
FT DISULFID 5 18 BY SIMILARITY.  
FT DISULFID 11 25 BY SIMILARITY.  
FT DISULFID 29 34 BY SIMILARITY.  
SQ SEQUENCE 250 AA; 27908 MW; 958E0816553C4985 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 250;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 VTQEFWD 97  
Db 50 VTQEFWD 56

RESULT 28  
ID SC65\_YEAST STANDARD; PRT; 273 AA.  
AC P29478;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Signal recognition particle SEC65 subunit.  
GN SEC65 OR YML105C OR YM8339.14C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92220169; PubMed=1313948;  
RA Stirling C.J., Hewitt E.W.;  
RT "The S. cerevisiae SEC65 gene encodes a component of yeast signal  
RT recognition particle with homology to human SRP19.";  
RL Nature 356:534-537(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C / AB972;  
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=92220168; PubMed=1313947;  
RA Hann B.C., Stirling C.J., Walter P.;  
RT "SEC65 gene product is a subunit of the yeast signal recognition  
RT particle required for its integrity.";  
RL Nature 356:532-533(1992).  
CC -1- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY HAS A CRUCIAL ROLE  
CC IN TARGETING SECRETORY PROTEINS TO THE ROUGH ENDOPLASMIC RETICULUM  
CC MEMBRANE. IT MUST BE INVOLVED INTIMATELY IN THE TRANSLLOCATION OF A  
CC WIDE VARIETY OF PROTEIN SUBSTRATES.  
CC -1- SUBUNIT: YEAST SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA  
CC MOLECULE (SCR1) AND AT LEAST SEVEN PROTEIN SUBUNITS: SRP72, SRP68,  
CC SRP54, SEC65, SRP21, SRP14 AND SRP7.  
CC -1- SIMILARITY: BELONGS TO THE SRP19 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X65783; CAA4666.1; -  
DR EMBL: Z49210; CAA89113.1; -  
DR PIR: S21731; S21731.  
DR SGD: S0004573; SEC65.  
DR InterPro: IPR002778; SRP19.  
DR Pfam: PF01922; SRP19; 1.  
DR ProDom: PD006609; SRP19; 1.

KW Signal recognition particle; RNA-binding; Ribonucleoprotein.  
SQ SEQUENCE 273 AA; 31170 MW; 4F4A6CB9225E385 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 273;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KDLEEVK 118  
Db 90 KDLEEVK 96

RESULT 29  
ID VAT1\_HUMAN STANDARD; PRT; 300 AA.  
AC Q99536; Q13035;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Synaptic vesicle membrane protein VAT-1 homolog.  
GN VAT1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97092865; PubMed=8938427;  
RA Smith T.M., Lee M.K., Szabo C.I., Jerome N., McEuen M., Taylor M.,  
RA Hood L., King M.-C.;  
RT "Complete genomic sequence and analysis of 117 kb of human DNA  
RT containing the gene BRCA1.";  
RL Genome Res. 6:1029-1049(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95293380; PubMed=7774926;  
RA Friedman L.S., Ostermeyer E.A., Lynch E.D., Szabo C.I., Meza J.E.,  
RA Anderson L.A., Dowd P., Lee M.K., Rowell S.E., Ellison J.,  
RA Boyd J., King M.-C.;  
RT "22 genes from chromosome 17q21: cloning, sequencing, and  
RT characterization of mutations in breast cancer families and tumors.";  
RL Genomics 25:256-263(1995).  
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE  
CC FAMILY. QUINONE OXIDOREDUCTASE SUBFAMILY. STRONG, TO TORPEDO  
CC CALIFORNICA VAT-1.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L78833; AAC37596.1; -  
DR EMBL: U18009; AAA95990.1; ALT\_INIT.  
DR MIM: 604631; -  
DR InterPro: IPR002085; Adh\_zn\_family.  
DR InterPro: IPR002364; QOR\_zeta\_crystal.  
DR Pfam: PF0107; adh\_zinc; 1.  
DR PROSITE: PS01162; QOR\_ZETA\_CRYSTAL; 1.  
KW Oxidoreductase; Zinc; Synapse; Membrane.  
FT CONFLICT 10 10 L -> R (IN REF. 2).  
FT CONFLICT 20 20 E -> M (IN REF. 2).  
FT CONFLICT 30 30 E -> G (IN REF. 2).  
FT CONFLICT 40 40 R -> D (IN REF. 2).  
FT CONFLICT 114 115 QL -> HV (IN REF. 2).  
FT CONFLICT 299 300 QN -> ES (IN REF. 2).  
SQ SEQUENCE 300 AA; 32558 MW; F81DDDBF31E8D7E4 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 300;

Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 203 FALKENG 209  
11111111  
Db 134 FALKENG 140

```

RESULT 30
RPOA_BACSU          STANDARD;      PRT;      314 AA.
ID  RPOA_BACSU
AC  P20429;
DT  01-FEB-1991 (Rel. 17, Created)
DT  01-FEB-1991 (Rel. 17, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (Transcriptase
DE  alpha chain) (RNA polymerase alpha subunit).
GN  RPOA.
OS  Bacillus subtilis.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group;
OC  Bacillus/Staphylococcus group; Bacillus.
OX  NCBI_TaxID=1423;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99213940; PubMed=2496109;
RA  Boylan S.A., Suh J.-W., Thomas S.M., Price C.W.;
RT  "Gene encoding the alpha core subunit of Bacillus subtilis RNA
RT  polymerase is cotranscribed with the genes for initiation factor 1
RT  and ribosomal proteins B, S13, S11, and L17."
RL  J. Bacteriol. 171:2553-2562(1989).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=168 / MARBURG;
RX  MEDLINE=96186897; PubMed=8635744;
RA  Suh J.-W., Boylan S.A., Oh S.H., Price C.W.;
RT  "Genetic and transcriptional organization of the Bacillus subtilis
RT  spc-alpha region."
RL  Gene 169:17-23(1996).
RN  [3]
RP  SEQUENCE OF 1-65 FROM N.A.
RX  MEDLINE=87008431; PubMed=3093467;
RA  Suh J.-W., Boylan S.A., Price C.W.;
RT  "Gene for the alpha subunit of Bacillus subtilis RNA polymerase maps
RT  in the ribosomal protein gene cluster."
RL  J. Bacteriol. 168:65-71(1986).
CC  -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC  OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC  SUBSTRATES.
CC  -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC  {RNA}(N).
CC  -!- SUBUNIT: CONSISTS OF A SIGMA FACTOR AND THE RNAP CORE ENZYME WHICH
CC  IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, 1 BETA' CHAIN AND 1
CC  OMEGA CHAIN.
CC  -!- DOMAIN: THE AMINO-TERMINAL PORTION IS INVOLVED IN THE ASSEMBLY OF
CC  CORE RNAP, WHEREAS THE C-TERMINAL IS INVOLVED IN INTERACTION WITH
CC  TRANSCRIPTIIONAL REGULATORS.
CC  -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE ALPHA CHAIN FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M26414; AAA22217.1; -
DR  EMBL; L47971; AAB06826.1; -
DR  EMBL; M13957; AAA22708.1; -
DR  EMBL; Z99104; CAB11919.1; -
DR  PIR; E32307; E32307.
DR  PIR; C24972; C24972.
DR  HSSP; P00574; 1BDF.

```

DR Subtilist; BG10732; rpoA.  
DR InterPro; IPR001700; RNA\_pol\_A\_bac.  
DR Pfam; PF01000; RNA\_pol\_A\_bac; 1.  
DR Pfam; PF03118; RNA\_pol\_A\_CTD; 1.  
DR ProDom; PD001179; RNA\_pol\_A\_bac; 1.  
KW Transiferase; Transcription; DNA-directed RNA polymerase;  
KW Complete proteome.  
SQ SEQUENCE 314 AA; 34799 MW; 9ADCAD891C0BCD67 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 LEEVAK 120  
11111111  
Db 296 LEEVAK 302

```

RESULT 31
Y034_TREPA          STANDARD;      PRT;      316 AA.
ID  Y034_TREPA
AC  O83077;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Putative periplasmic metal-binding protein TP0034 precursor.
GN  TP0034.
OS  Treponema pallidum.
OC  Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX  NCBI_TaxID=160;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NICHOLS;
RX  MEDLINE=98332770; PubMed=9665876;
RA  Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA  Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA  Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA  Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA  McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA  Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA  Venter J.C.;
RT  "Complete genome sequence of Treponema pallidum, the syphilis
RT  spirochete."
RL  Science 281:375-388(1998).
CC  -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC  TP0034/TP0035/TP0036 FOR A METAL. METAL-BINDING COMPONENT.
CC  -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC  -!- SIMILARITY: BELONGS TO THE BACTERIAL SOLUTE-BINDING PROTEIN FAMILY
CC  9.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE001188; AAC65029.1; -
DR  TIGR; TP0034; -
DR  InterPro; IPR001987; Lipoprotein_4.
DR  Pfam; PF01297; Lipoprotein_4; 1.
KW Hypothetical protein; Transport; Periplasmic; Metal-binding; Signal;
KW Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 316 PUTATIVE PERIPLASMIC METAL-BINDING
FT FT PROTEIN TP0034.
SQ SEQUENCE 316 AA; 35433 MW; 16051C2199BC81AB CRC64;

```

Query Match 2.6%; Score 7; DB 1; Length 316;  
Best Local Similarity 100.0%; Pred. No. 44;



Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205  
 11111111  
 Db 184 AARLEAL 190

# RESULT 32 EXL2\_HUMAN

ID EXL2\_HUMAN STANDARD; PRT; 330 AA.  
 AC Q9UBQ6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Exostosin-like 2 (EXT-related protein 2).  
 GN EXTL2 OR EXTR2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=98111853; PubMed=9450183;  
 RA Wuyts W., Van Hul W., Hendrickx J., Speleman F., Wauters J.,  
 De Boule K., Van Roy N., Van Agtmael T., Bossuyt P., Willems P.J.;  
 RT "Identification and characterization of a novel member of the EXT gene  
 family, EXTL2.";  
 RL Eur. J. Hum. Genet. 5:382-389(1997).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=98139867; PubMed=9473480;  
 RA Saito T., Seki N., Yamauchi M., Tsuji S., Hayashi A., Kozuma S.,  
 Hori T.-A.;

RT "Structure, chromosome location, and expression profile of EXTR1 and  
 EXTR2, new members of the multiple exostosin gene family.";  
 RL Biochem. Biophys. Res. Commun. 243:61-66(1998).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 reticulum (By similarity).

CC -1- TISSUE SPECIFICITY: UBIQUITOUS.

CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF000416; AAC02898.1; -.  
 DR EMBL; AB009284; BAA24081.1; -.  
 DR MIM; 602411; -.  
 DR InterPro; IPR004263; Exostosin.  
 DR Pfam; PF03016; Exostosin; 1.  
 KW Multigene family; Transmembrane; Signal-anchor.  
 FT TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 SQ CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SEQUENCE 330 AA; 37465 MW; 6976BE7EC6F588C8 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 DNLEKET 103  
 11111111  
 Db 270 DNLEKET 276

RESULT 33  
 DIA2\_MOUSE STANDARD; PRT; 349 AA.  
 ID DIA2\_MOUSE

AC 070566;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Diaphanous protein homolog 2 (Diaphanous-related formin 2) (DRF2)  
 DE (mdia3) (Fragment).  
 GN DIAPH2 OR DIAP2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=98163437; PubMed=9497258;  
 RA Bione S., Sala C., Manzini C., Arrigo G., Zuffardi O., Banfi S.,  
 Borsani G., Jonveaux P., Philippe C., Zuccotti M., Ballabio A.,  
 Toniolo D.;  
 RT "A human homologue of the Drosophila melanogaster diaphanous gene is  
 disrupted in a patient with premature ovarian failure: evidence for  
 conserved function in oogenesis and implications for human  
 sterility.";  
 RL Am. J. Hum. Genet. 62:533-541(1998).

CC -1- FUNCTION: MAY BE INVOLVED IN OOGENESIS.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN LIVER, HEART, KIDNEY, OVARY AND  
 TESTIS, AT E16, P6 AND P16.

CC -1- DOMAIN: DRFS ARE REGULATED BY INTRAMOLECULAR GBD-DAD BINDING WHERE  
 RHO-GTP ACTIVATES THE DRFS BY DISRUPTING THE GBD-DAD INTERACTION  
 (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 2 (FH2) DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 DRF AUTOREGULATORY DOMAIN (DAD).

CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOLOG FAMILY. DIAPHANOUS  
 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Y15910; CAA75871.1; -.  
 DR MGD; MGI:1858500; Diap2.  
 DR InterPro; IPR003104; FH2.  
 DR Pfam; PF02181; FH2; 1.  
 DR SMART; SM00498; FH2; 1.  
 KW Coiled coil.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 319 FH2.  
 FT DOMAIN 250 301 COILED COIL (POTENTIAL).  
 FT DOMAIN 302 316 DAD.  
 FT DOMAIN 320 323 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 286 289 POLY-LYS.  
 SQ SEQUENCE 349 AA; 40557 MW; B79C15CE0E4F758 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 349;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 LEALKEN 208  
 11111111  
 Db 251 LEALKEN 257

RESULT 34  
 REPB\_THIFE STANDARD; PRT; 352 AA.  
 ID REPB\_THIFE  
 AC P27186;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)



DE DNA primase (EC 2.7.7.-) (Replication primase).  
GN REPB.  
OS Thiobacillus ferrooxidans.  
OG Plasmid pTF-FC2.  
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
OX NCBI\_TaxID=920;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92104508; PubMed=1761233;  
RA Dorrington R.A., Bardien S., Rawlings D.E.;  
RT "The broad-host-range plasmid pTF-FC2 requires a primase-like protein  
for autonomous replication in Escherichia coli.";  
RL Gene 108:7-14(1991).  
CC -1- FUNCTION: FUNCTIONS AS A PRIMASE WITH RESPECT TO REPLICATION AT  
CC THE ORIGIN OF REPLICATION (VEGETATIVE) OF PTF-FC2.  
CC -1- SIMILARITY: SOME, TO REPB' PRIMASE OF PLASMID INCQ RSF1010.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M64981; AAA27386.1; -.  
DR PIR; JS0637; JS0637.  
KW DNA replication; Transferrase; DNA-directed RNA polymerase; Plasmid.  
SQ SEQUENCE 352 AA; 40156 MW; 6B5E63D06FA59E24 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 352;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 KAKPALE 236  
|||||||  
Db 216 KAKPALE 222

RESULT 35  
C11D\_MOUSE  
ID C11D\_MOUSE STANDARD; PRT; 359 AA.  
AC Q9DD19;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein c11orf13 homolog.  
GN C11ORF13.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordong P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch X.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AK02243; BAB21960.1; -.  
DR EMBL; BC01131; AAH1131.1; -.  
DR InterPro; IPR000159; RA.  
DR SMART; SM00314; RA; 1.  
KW Coiled coil.  
FT DOMAIN 135 154 PRO-RICH.  
FT DOMAIN 180 208 COILED COIL (POTENTIAL).  
FT DOMAIN 242 301 COILED COIL (POTENTIAL).  
FT DOMAIN 309 328 PRO-RICH.  
FT CONFLICT 53 53 L -> P (IN REF. 1).  
FT CONFLICT 350 350 V -> F (IN REF. 1).  
FT CONFLICT 353 359 MISSING (IN REF. 1).  
SQ SEQUENCE 359 AA; 39246 MW; 39EF0BCA1D8500F1 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 199 AARLEAL 205  
|||||||  
Db 209 AARLEAL 215

RESULT 36  
P53\_PLAEE  
ID P53\_PLAEE STANDARD; PRT; 366 AA.  
AC O12946;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cellular tumor antigen p53 (Tumor suppressor p53).  
GN TP53 OR P53.  
OS Platycthyes flesus (European flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Pleuronectidae; Platycthyes.  
OX NCBI\_TaxID=8260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99138245; PubMed=9972298;  
RA Cachot J., Galgani F., Vincent F.;  
RT "cDNA cloning and expression analysis of flounder p53 tumour  
suppressor gene.";  
RL Comp. Biochem. Physiol. 121B:235-242(1998).  
CC -1- FUNCTION: ACTS AS A TUMOR SUPPRESSOR IN MANY TUMOR TYPES. INDUCES  
CC GROWTH ARREST OR APOPTOSIS DEPENDING ON THE PHYSIOLOGICAL  
CC CIRCUMSTANCES OR CELL TYPE, BUT BOTH ACTIVITIES ARE INVOLVED IN  
CC TUMOR SUPPRESSION. ACTS IN CELL CYCLE REGULATION, IS A TRANS-  
CC ACTIVATOR THAT ACTS TO NEGATIVELY REGULATE CELL DIVISION BY  
CC CONTROLLING A SET OF GENES REQUIRED FOR THIS PROCESS. ONE OF THE  
CC ACTIVATED GENES IS AN INHIBITOR OF CYCLIN-DEPENDENT KINASES.  
CC APOPTOSIS INDUCTION SEEMS TO BE MEDIATED EITHER BY STIMULATION OF  
CC BAX AND FAS ANTIGEN EXPRESSION, OR BY REPRESSION OF BCL-2

```
CC      EXPRESSION (BY SIMILARITY).
CC      -1- SUBUNIT: BINDS DNA AS AN HOMOTETRAMER (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Nuclear.
CC      -1- SIMILARITY: BELONGS TO THE P53 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y08919; CAA70123.1; -.
DR      HSSP; P04637; 1YCS.
DR      InterPro; IPR002117; P53.
DR      Pfam; PF00870; P53; 1.
DR      PRINTS; PR00386; P53SUPPRESSR.
DR      ProDom; PD002681; P53; 1.
DR      PROSITE; PS00348; P53; 1.
KW      Anti-oncogene; DNA-binding; Transcription regulation; Activator;
KW      Nuclear protein; Phosphorylation; Apoptosis.
FT      DOMAIN 1 41 TRANSCRIPTION ACTIVATION (ACIDIC).
FT      DNA_BIND 80 267 BY SIMILARITY.
FT      DOMAIN 305 336 OLIGOMERIZATION.
FT      DOMAIN 341 362 BASIC (REPRESSION OF DNA-BINDING).
FT      DOMAIN 282 298 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT      MOD_RES 365 365 PHOSPHORYLATION (BY SIMILARITY).
SQ      SEQUENCE 366 AA; 40619 MW; 40DE4CA20B22BB88 CRC64;

Query Match      2.6%; Score 7; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      76 SVTSTFS 82
      |||||
Db      99 SVTSTFS 105

RESULT 37
YCCZ_ECOLI STANDARD; PRT; 379 AA.
AC      P75881;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Putative polysaccharide export protein yccz precursor.
GN      YCCZ OR B0983 OR Z1400 OR ECS1139.
OS      Escherichia coli, and
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562, 83334;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12;
RX      MEDLINE=97061202; PubMed=8905232;
RA      Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA      Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA      Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA      Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA      Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
```

```
RA      Yano M., Horiuchi T.;
RT      "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT      corresponding to the 12.7-28.0 min region on the linkage map.";
RL      DNA Res. 3:137-155(1996).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blattner F.R.;
RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL      Nature 409:529-533(2001).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / RIMD 0509952;
RX      MEDLINE=21156231; PubMed=11258796;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA      Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT      "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT      O157:H7 and genomic comparison with a laboratory strain K-12.";
RL      DNA Res. 8:11-22(2001).
CC      -1- FUNCTION: MAY BE INVOLVED IN POLYSACCHARIDE TRANSPORT.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
CC      (By similarity).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE000200; AAC74068.1; -.
DR      EMBL; D90735; BAA35748.1; -.
DR      EMBL; D90736; BAA36123.1; -.
DR      EMBL; AE005292; AAG55531.1; -.
DR      EMBL; AP002554; BAB34562.1; -.
DR      EcoGene; EG13728; YCCZ.
DR      InterPro; IPR003715; Poly-export.
DR      Pfam; PF02563; Poly-export; 1.
DR      PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW      Hypothetical protein; Polysaccharide transport; Transport;
KW      Outer membrane; Transmembrane; Lipoprotein; Porin; Signal;
KW      Complete proteome.
FT      SIGNAL 1 20 POTENTIAL.
FT      CHAIN 21 379 PUTATIVE POLYSACCHARIDE EXPORT PROTEIN
FT      LIPID 21 21 YCCZ.
FT      SEQUENCE 379 AA; 41740 MW; 36A5F5B9E9C5E842 CRC64;

Query Match      2.6%; Score 7; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 VLTAVL 11
      |||||
Db      10 VLTAVL 16

RESULT 38
DXR_NEIMA STANDARD; PRT; 394 AA.
ID      DXR_NEIMA
AC      Q9JX33;
DT      16-OCT-2001 (Rel. 40, Created)
```

```

DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE 1-deoxy-D-xylose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
DE reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR NMA0083.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT *Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -1- FUNCTION: CATALYZES THE NADP-DEPENDENT REARRANGEMENT AND REDUCTION
CC OF 1-DEOXY-D-XYULOSE-5-PHOSPHATE (DXP) TO 2-C-METHYL-D-ERYTHRITOL
CC 4-PHOSPHATE (MEP) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xyulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL162752; CAB83399.1; -
DR InterPro: IPR003821; DXP_reductoisomerase.
DR Pfam: PF02670; DXP_reductoisom; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 9 16 NADPH (POTENTIAL).
SQ SEQUENCE 394 AA; 41912 MW; B5D361B224806D6B CRC64;

Query Match 2.6%; Score 7; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205
Db 64 AARLEAL 70

RESULT 39
DXR_NEIMB
ID DXR_NEIMB STANDARD; PRT; 394 AA.
AC O9K1G8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE 1-deoxy-D-xyulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
DE reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR NMB0184.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RT tetrilin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

```

```

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciltone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: CATALYZES THE NADP-DEPENDENT REARRANGEMENT AND REDUCTION
CC OF 1-DEOXY-D-XYLULOSE-5-PHOSPHATE (DXP) TO 2-C-METHYL-D-ERYTHRITOL
CC 4-PHOSPHATE (MEP) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE002375; AAF40641.1; -.
CC TIGR: NMB0184; -.
DR InterPro: IPR003821; DXP_reductoisomerase.
DR Pfam: PF02670; DXP_reductoisom; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 9 16 NADPH (POTENTIAL).
SQ SEQUENCE 394 AA; 41920 MW; AA85397E5BE7BD2E CRC64;

Query Match 2.6%; Score 7; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AARLEAL 205
Db 64 AARLEAL 70

RESULT 40
HYFL_ALCEU
ID HYFL_ALCEU STANDARD; PRT; 394 AA.
AC P45805;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Hydrogenase maturation protein hypF1.
GN HYPF1 OR HYPF.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Plasmid megaplasmid pHG1.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=H16 / ATCC 17699;
RX MEDLINE=93356597; PubMed=8352644;
RA Dornedde J., Eitinger M., Friedrich B.;
RT "Analysis of a pleiotropic gene region involved in formation of
RT catalytically active hydrogenases in Alcaligenes eutrophus H16.";
RL Arch. Microbiol. 159:545-553(1993).
RN [2]
RP REVISIONS.
RC STRAIN=H16 / ATCC 17699;
RA Eitinger T.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE HYDROGENASE MATURATION PROCESS.

```



CC -!- SIMILARITY: BELONGS TO THE HYF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.ch](mailto:license@isb.ch)).  
 CC -----  
 DR EMBL; X70183; CAA49731.1; -  
 DR HSSP; P02647; IODR.  
 KW plasmid.  
 SQ SEQUENCE 394 AA; 40677 MW; 268E9CF0C9B3E64B CRC64;  
 QY 196 ORLAARL 202  
 Db 85 ORLAARL 91  
 Query Match 2.6%; Score 7; DB 1; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 41  
 ID APA4\_HUMAN STANDARD; PRT; 396 AA.  
 AC P06727;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Apolipoprotein A-IV precursor (Apo-AIV).  
 GN APOA4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Intestine;  
 RX MEDLINE=89194198; PubMed=2930771;  
 RA Yang C., Gu Z.W., Xiong W., Rosseneu M., Yang H.X., Lee B.M.,  
 RA Gotto A.M. Jr., Chan L.;  
 RT "The primary structure of human apolipoprotein A-IV.";  
 RL Biochim. Biophys. Acta 1002:231-237(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87041474; PubMed=3095836;  
 RA Karathanasis S.K., Oeltgen P., Haddad I.A., Antonarakis S.E.;  
 RT "Structure, evolution, and polymorphisms of the human apolipoprotein  
 RL A4 gene (APOA4).";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86296629; PubMed=3755616;  
 RA Karathanasis S.K., Yunis I.;  
 RT "Structure, evolution, and tissue-specific synthesis of human  
 RL apolipoprotein AIV";  
 RL Biochemistry 25:3962-3970(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87250378; PubMed=3036793;  
 RA Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,  
 RA Gordon J.I., Taylor J.M.;  
 RT "Structure and expression of the human apolipoprotein A-IV gene.";  
 RL J. Biol. Chem. 262:7973-7981(1987).  
 RN [5]  
 RP SEQUENCE OF 21-396 FROM N.A.  
 RX MEDLINE=86111885; PubMed=3080432;  
 RA Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;  
 RT "The nucleotide and derived amino acid sequence of human  
 RT apolipoprotein A-IV mRNA and the close linkage of its gene to the

RT genes of apolipoproteins A-I and C-III.";  
 RL J. Biol. Chem. 261:1998-2002(1986).  
 RN [6]  
 RP SIGNAL SEQUENCE CLEAVAGE SITE.  
 RX MEDLINE=84161950; PubMed=6706947;  
 RA Gordon J.I., Bisglier C.L., Sims H.F., Sachdev O.P., Glickman R.M.,  
 RA Strauss A.W.;  
 RT "Biosynthesis of human preapolipoprotein A-IV.";  
 RL J. Biol. Chem. 259:468-474(1984).  
 RN [7]  
 RP REVIEW ON POLYMORPHISM.  
 RA Lohse P., Brewer H.B. Jr.;  
 RT "Genetic polymorphism of apolipoprotein A-IV.";  
 RL Curr. Opin. Lipidol. 2:90-95(1991).  
 RN [8]  
 RP VARIANT A-IV\*2.  
 RX MEDLINE=90277616; PubMed=2351649;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to  
 RT nucleotide substitutions in the apolipoprotein A-IV gene.";  
 RL J. Biol. Chem. 265:10061-10064(1990).  
 RN [9]  
 RP VARIANTS A-IV\*0 AND A-IV\*3.  
 RX MEDLINE=90324273; PubMed=1973689;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for  
 RT two rare variants of apolipoprotein A-IV-1.";  
 RL J. Biol. Chem. 265:12734-12739(1990).  
 RN [10]  
 RP VARIANTS.  
 RX MEDLINE=91310615; PubMed=1677358;  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
 RT 1(Thr-347->Ser), apoA-IV-0(Lys-167->Glu, Gln-360->His), and apoA-IV-  
 RT 3(Glu-165->Lys).";  
 RL J. Biol. Chem. 266:13513-13518(1991).  
 RN [11]  
 RP ERRATUM.  
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RL J. Biol. Chem. 266:19866-19866(1991).  
 RN [12]  
 RP VARIANT MET-13.  
 RX MEDLINE=92238494; PubMed=1349197;  
 RA von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
 RA Assmann G.;  
 RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV  
 RT gene are associated with changes in the concentration of apo B- and  
 RL apo A-I-containing lipoproteins in a normal population.";  
 RL Am. J. Hum. Genet. 50:1115-1128(1992).  
 RN [13]  
 RP VARIANT SER-147.  
 RX MEDLINE=92144647; PubMed=1737067;  
 RA Tenkanen H., Koskinen P., Metso J., Baumann M., Lukka M.,  
 RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
 RA Manninen V., Ehnholm C.;  
 RT "A novel polymorphism of apolipoprotein A-IV is the result of an  
 RT asparagine to serine substitution at residue 127.";  
 RL Biochim. Biophys. Acta 1138:27-33(1992).  
 RN [14]  
 RP VARIANT A-IV\*5.  
 RX MEDLINE=93138374; PubMed=1487136;  
 RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
 RA Ferrell R.E., Pollitzer W.S.;  
 RT "Molecular basis of a unique African variant (A-IV 5) of human  
 RT apolipoprotein A-IV and its significance in lipid metabolism.";  
 RL Genet. Epidemiol. 9:379-388(1992).  
 RN [15]  
 RP VARIANTS LYS-44 (BUDAPEST-2); CYS-305 AND SER-367 (BUDAPEST-1).  
 RX MEDLINE=95245341; PubMed=7728150;  
 RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
 RA Csaszar A.;  
 RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene  
 RT frequencies, effect on lipid levels, and sequence of two new

```

RT variants."
RL Hum. Mutat. 5:58-65(1995).
RN [16]
RP VARIANTS FCHL SEATTLE SER-161; LEU-178 AND GLN-264.
RX MEDLINE-97114287; Pubmed-8956036;
RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;
RT "Two novel apolipoprotein A-IV variants in individuals with familial
combined hyperlipidemia and diminished levels of lipoprotein lipase
activity."
RL Hum. Mutat. 8:319-325(1996).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.
CC SECRETED IN PLASMA.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV*0
CC TO APOA-IV*7). APOA-IV*1 IS THE MAJOR ALLELE (90%), IV*2 IS ALSO
CC COMMON (8%), THE OTHERS ARE RARE ALLELES.
CC -1- DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN
CC OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF
CC FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14642; AAA51745.1; -
DR EMBL; X13629; CAA31955.1; -
DR EMBL; M14566; AAA51748.1; -
DR EMBL; J02758; AAA96731.1; -
DR EMBL; M13654; AAA51744.1; -
DR PIR; A26481; LPHUA4.
DR PIR; A24449; A24449.
DR PIR; A29330; A29330.
DR PIR; A26280; A26280.
DR PIR; S02715; S02715.
DR HSSP; P02649; INFO.
DR SWISS-2DPAGE; P06727; HUMAN.
DR MIM; 107690; -
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;
KW Disease mutation; Polymorphism.
FT SIGNAL 1 20
FT CHAIN 1 396
FT DOMAIN 33 330
FT REPEAT 33 54
FT REPEAT 60 81
FT REPEAT 82 103
FT REPEAT 115 136
FT REPEAT 137 158
FT REPEAT 159 180
FT REPEAT 181 202
FT REPEAT 203 224
FT REPEAT 225 246
FT REPEAT 247 268
FT REPEAT 269 286
FT REPEAT 287 308
FT REPEAT 309 330
FT DOMAIN 372 389
FT GLU/GLN-RICH.

```

```

FT VARIANT 13 13 V -> M (IN APOA-IV*1D).
FT FT /FTid=VAR_000626.
FT VARIANT 44 44 E -> K (IN BUDPEST-2).
FT FT /FTid=VAR_000627.
FT VARIANT 147 147 N -> S (IN APOA-IV*1B).
FT FT /FTid=VAR_000628.
FT VARIANT 161 161 A -> S (IN SEATTLE-3; IN FCHL).
FT FT /FTid=VAR_000629.

Query Match 2.6%; Score 7; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ELRQRLA 199
Db 262 ELRQRLA 268

RESULT 42
AP04_PAPAN STANDARD; PRT; 401 AA.
ID AP04_PAPAN
AC 028758;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV) (Fragment).
GN APOA4.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_Taxid=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=93340170; Pubmed=8101842;
RA Hixson J.E., Kammerer C.M., Mott G.E., Britten M.L., Birnbaum S.,
RA Powers P.K., Vandenberg J.L.;
RT "Baboon apolipoprotein A-IV. Identification of Lys-76-->Glu that
RT distinguishes two common isoforms and detection of length
RT polymorphisms at the carboxyl terminus.";
RL J. Biol. Chem. 268:15667-15673(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.
CC SECRETED IN PLASMA.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- POLYMORPHISM: THERE ARE TWO COMMON APOA-IV ISOFORMS, I (SHOWN
CC HERE) AND E. THE I ISOFORM IS ASSOCIATED WITH HIGHER LEVELS OF
CC HIGH DENSITY LIPOPROTEIN-C ON A HIGH CHOLESTEROL, SATURATED FAT
CC DIET.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13174; AAA35379.1; -
DR HSSP; P02649; INFO.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.

```



KW Plasma; lipid transport; HDL; Chylomicron; Repeat; Signal;  
KW Polymorphism.  
FT NON\_TER 1 1  
FT SIGNAL <1 4 POTENTIAL.  
FT CHAIN 5 401 APOLIPROTEIN A-IV.  
FT DOMAIN 17 314 13 X 22 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 17 38 1.  
FT REPEAT 44 65 2.  
FT REPEAT 66 87 3.  
FT REPEAT 99 120 4.  
FT REPEAT 121 142 5.  
FT REPEAT 143 164 6.  
FT REPEAT 165 186 7.  
FT REPEAT 187 208 8.  
FT REPEAT 209 230 9.  
FT REPEAT 231 252 10.  
FT REPEAT 253 270 11.  
FT REPEAT 271 292 12.  
FT REPEAT 293 314 13.  
FT DOMAIN 356 394 GLU/GLN-RICH.  
FT VARIANT 80 80 K -> E (IN ISOFORM E).  
SQ SEQUENCE 401 AA; 46538 MW; 0A76D1284AA9837F CRC64;

Query Match 2.6%; Score 7; DB 1; Length 401;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 ELRQRLA 199  
DB 246 ELRQRLA 252

RESULT 43  
AATC\_BOVIN  
ID AATC\_BOVIN STANDARD; PRT; 412 AA.  
AC P33097;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)  
DE (Glutamate oxaloacetate transaminase-1).  
GN GOT1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Heart;  
RA MEDLINE=94039977; PubMed=8224363;  
RA Aurilia V., Palmisano A., Ferrara L., Cubellis M.V., Sanna G.,  
RA Marino G.;  
RT "Cloning and sequence analysis of A cDNA encoding bovine cytosolic  
RT aspartate aminotransferase.";  
RL Int. J. Biochem. 25:1505-1509(1993).  
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +  
CC L-glutamate.  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC  
CC ONE AND A MITOCHONDRIAL ONE.  
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; X66020; CAA46818.1; -.  
DR PIR; S21560; S21560.  
DR HSSP; P00503; 1AJS.  
DR InterPro; IPR001511; Aminotran\_1.  
DR InterPro; IPR000796; Aminotransfrse\_sub.  
DR Pfam; PF00155; aminotran\_1\_2; 1.  
DR PRINTS; PR00799; TRANSAMINASE.  
DR PROSITE; PS00105; AA\_TRANSFR\_CLASS\_1; 1.  
KW Transferase; Aminotransferase; Pyridoxal phosphate.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT BINDING 258 258 PYRIDOXAL PHOSPHATE.  
SQ SEQUENCE 412 AA; 46281 MW; 1F7DACCCID7FBC62 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 412;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 ARLEALK 206  
DB 339 ARLEALK 345

RESULT 44  
AATC\_HORSE  
ID AATC\_HORSE STANDARD; PRT; 412 AA.  
AC P08906;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)  
DE (Glutamate oxaloacetate transaminase-1).  
GN GOT1.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=87169770; PubMed=3104605;  
RA Doonan S., Martini F., Angelaccio S., Pascarella S., Barra D.,  
RA Bossa F.;  
RT "The complete amino acid sequences of cytosolic and mitochondrial  
RT aspartate aminotransferases from horse heart, and inferences on  
RT evolution of the isoenzymes.";  
RL J. Mol. Evol. 23:328-335(1986).  
RN [2]  
RP PARTIAL SEQUENCE.  
RX MEDLINE=84281020; PubMed=6466688;  
RA Martini F., Angelaccio S., Barra D., Doonan S., Bossa F.;  
RT "Partial amino-acid sequence and cysteine reactivities of cytosolic  
RT aspartate aminotransferase from horse heart.";  
RL Biochim. Biophys. Acta 789:51-56(1984).  
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +  
CC L-glutamate.  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC  
CC ONE AND A MITOCHONDRIAL ONE.  
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES.  
CC PIR; A26341; A26341.  
DR PIR; A05067; A05067.  
DR HSSP; P00503; 1AJS.  
DR InterPro; IPR001511; Aminotran\_1.  
DR InterPro; IPR000796; Aminotransfrse\_sub.  
DR Pfam; PF00155; aminotran\_1\_2; 1.  
DR PRINTS; PR00799; TRANSAMINASE.  
DR PROSITE; PS00105; AA\_TRANSFR\_CLASS\_1; 1.  
KW Transferase; Aminotransferase; Pyridoxal phosphate; Acetylation.

FT MOD\_RES 1 1 ACETYLATION.  
 FT BINDING 258 258 PYRIDOXAL PHOSPHATE.  
 SQ SEQUENCE 412 AA; 46213 MW; 2997578584920D3D CRC64;

Query Match 2.6%; Score 7; DB 1; Length 412;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 ARLEALK 206  
 |||||  
 Db 339 ARLEALK 345

RESULT 45  
 AATC\_HUMAN STANDARD; PRT; 412 AA.  
 ID AATC\_HUMAN  
 AC P17174;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)  
 DE (Glutamate oxaloacetate transaminase-1).  
 GN GOT1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE.  
 RC TISSUE=Liver;  
 RX MEDLINE=91054353; PubMed=2241899;  
 RA Doyle J.M., Schinina M.E., Bossa F., Doonan S.;  
 RT "The amino acid sequence of cytosolic aspartate aminotransferase from  
 RT human liver.";  
 RL Biochem. J. 270:651-657(1990).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=90344765; PubMed=1974457;  
 RA Bousquet-Lemercler B., Pol S., Pave-Preux M., Hanoune J., Barouki R.;  
 RT "Properties of human liver cytosolic aspartate aminotransferase mRNAs  
 RT generated by alternative polyadenylation site selection.";  
 RL Biochemistry 29:5293-5299(1990).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA Wang C.Y., Huang Y.Q., Shi J.D., Marron M.P., Ruan Q.G.,  
 RA Hawkins-Lee B., Ochoa B., She J.X.;  
 RT "Genomic structure and mutation analysis of GOT1 in the urofacial  
 RT (Ochoa) syndrome gene critical region on chromosome 10.";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Yu W., Sarginson J., Gibbs R.A.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +  
 CC L-glutamate.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC  
 CC ONE AND A MITOCHONDRIAL ONE.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M37400; AAA35563.1; -  
 DR EMBL; AF080467; AAC32851.1; -  
 DR EMBL; AF080459; AAC32851.1; JOINED.  
 DR EMBL; AF080460; AAC32851.1; JOINED.  
 DR EMBL; AF080461; AAC32851.1; JOINED.  
 DR EMBL; AF080462; AAC32851.1; JOINED.  
 DR EMBL; AF080463; AAC32851.1; JOINED.  
 DR EMBL; AF080464; AAC32851.1; JOINED.  
 DR EMBL; AF080465; AAC32851.1; JOINED.  
 DR EMBL; AF080466; AAC32851.1; JOINED.  
 DR EMBL; AF052153; AAC28622.1; -  
 DR PIR; S13035; S13035.  
 DR HSSP; P00503; IAJ5.  
 DR MIM; 138180; -  
 DR InterPro; IPR001511; Aminotran\_1.  
 DR InterPro; IPR000796; Aminotransfrse\_sub.  
 DR Pfam; PF00155; aminotran\_1\_2; 1.  
 DR PRINTS; PR00799; TRANSAMINASE.  
 DR PROSITE; PS00105; AA\_TRANSFR\_CLASS\_1; 1.  
 KW Transferase; Aminotransferase; Pyridoxal phosphate.  
 FT INIT\_MET 0 0  
 FT BINDING 258 258 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT CONFLICT 214 214 H -> R (IN REF. 1).  
 FT SEQUENCE 412 AA; 46116 MW; 9863F47FBCD055AF CRC64;

Query Match 2.6%; Score 7; DB 1; Length 412;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 ARLEALK 206  
 |||||  
 Db 339 ARLEALK 345

RESULT 46  
 AATC\_MOUSE STANDARD; PRT; 412 AA.  
 ID AATC\_MOUSE  
 AC P05201;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)  
 DE (Glutamate oxaloacetate transaminase-1).  
 GN GOT1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87057413; PubMed=3782150;  
 RA Obaru K., Nomiya H., Shimada K., Nagashima F., Morino Y.;  
 RT "Cloning and sequence analysis of mRNA for mouse aspartate  
 RT aminotransferase isoenzymes.";  
 RL J. Biol. Chem. 261:16976-16983(1986).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88245178; PubMed=3379636;  
 RA Obaru K., Tsuzuki T., Setoyama C., Shimada K.;  
 RT "Structural organization of the mouse aspartate aminotransferase  
 RT isoenzyme genes. Introns antedate the divergence of cytosolic and  
 RT mitochondrial isoenzyme genes.";  
 RL J. Mol. Biol. 200:13-22(1988).  
 CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +  
 CC L-glutamate.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC  
 CC ONE AND A MITOCHONDRIAL ONE.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT

```

CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02623; AAA37263.1; .
DR EMBL; X07302; CAA30275.1; .
DR EMBL; X07303; CAA30275.1; JOINED.
DR EMBL; X07304; CAA30275.1; JOINED.
DR EMBL; X07305; CAA30275.1; JOINED.
DR EMBL; X07306; CAA30275.1; JOINED.
DR EMBL; X07307; CAA30275.1; JOINED.
DR EMBL; X07308; CAA30275.1; JOINED.
DR EMBL; X07309; CAA30275.1; JOINED.
DR PIR; S01076; S01076.
DR PIR; B25349; B25349.
DR HSSP; P00503; IAJ5.
DR SWISS-2DPAGE; P05201; MOUSE.
DR MGD; MGI:95791; Got1.
DR InterPro; IPR001511; Aminotran_1.
DR InterPro; IPR000796; Aminotransfrse_sub.
DR Pfam; PF00155; aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
DR Transferase; Aminotransferase; Pyridoxal phosphate.
FT INIT_MET 0
FT BINDING 258 258 PYRIDOXAL PHOSPHATE.
FT CONFLICT 290 290 I -> N (IN REF. 2).
SQ SEQUENCE 412 AA; 46100 MW; 97EB1FED83E6AA23 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 ARLEALK 206
Db 339 ARLEALK 345

RESULT 47
AATC_PIG STANDARD; PRT; 412 AA.
AC P00503;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)
DE (Glutamate oxaloacetate transaminase-1).
GN GOT1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89229121; PubMed=2653435;
RA Nagashima F., Tanase S., Fukumoto Y., Joh T., Nomiya H.,
RA Tsuzuki T., Shimada K., Kuramitsu S., Kagamiyama H., Morino Y.;
RT "cDNA cloning and expression of pig cytosolic aspartate
RT aminotransferase in Escherichia coli: amino-terminal heterogeneity of
RT expressed products and lack of its correlation with enzyme
RT function.";
RL Biochemistry 28:1153-1160(1989).
RN [2]
RP SEQUENCE.
RC TISSUE=Heart muscle;
RA Ovchinnikov Y.A., Egorov T.A., Aldanova N.A., Feigina M.Y.,

```

```

RA Lipkin V.M., Abdulaev N.G., Grishin E.V., Kiselev A.P.,
RA Modyanov N.N., Braunstein A.E., Polyanovsky O.L., Nosikov V.V.;
RT "The complete amino acid sequence of cytoplasmic aspartate
RT aminotransferase from pig heart.";
RL FEBS Lett. 29:31-34(1973).
RN [3]
RP EXPERIMENTAL DETAILS.
RA Ovchinnikov Y.A., Egorov T.A., Aldanova N.A., Feigina M.Y.,
RA Lipkin V.M., Abdulaev N.G., Grishin E.V., Kiselev A.P.,
RA Modyanov N.N., Braunstein A.E., Polyanovsky O.L., Nosikov V.V.;
RT "The complete primary structure of cytoplasmic aspartate amino-
RT transferase from pig heart muscle.";
RL Izv. Akad. Nauk SSSR, Ser. Khim. 1974:1189-1196(1974).
RN [4]
RP SEQUENCE.
RX MEDLINE=76087776; PubMed=1239277;
RA Doonan S., Doonan H.J., Hanford R., Vernon C.A., Walker J.M.,
RA Airolidi L.P., Da S., Bossa F., Barra D., Carloni M., Fasella P.,
RA Riva F.;
RT "The primary structure of aspartate aminotransferase from pig heart
RT muscle. Digestion with a proteinase having specificity for lysine
RT residues.";
RL Biochem. J. 149:497-506(1975).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=73044407; PubMed=4634443;
RA Polyanovsky O.L., Demidkina T.V., Egorov C.A.;
RT "The position of an essential tyrosine residue in the polypeptide
RT chain of aspartate transaminase.";
RL FEBS Lett. 23:262-264(1972).
RN [6]
RP ACTIVE SITE.
RX MEDLINE=69285398; PubMed=5809231;
RA Morino Y., Watanabe T.;
RT "Primary structure of pyridoxal phosphate binding site in the
RT mitochondrial and extramitochondrial aspartate aminotransferases from
RT pig heart muscle. Chymotryptic peptides.";
RL Biochemistry 8:3412-3417(1969).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.74 ANGSTROMS).
RX MEDLINE=97362209; PubMed=9211866;
RA Rhee S., Silva M.M., Hyde C.C., Rogers P.H., Metzler C.M.,
RA Metzler D.E., Arnone A.;
RT "Refinement and comparisons of the crystal structures of pig
RT cytosolic aspartate aminotransferase and its complex with
RT 2-methylaspartate.";
RL J. Biol. Chem. 272:17293-17302(1997).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -1- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/A/GOT.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24088; AAA53531.1; .
DR PIR; A00592; XNPGDC.
DR PIR; A30138; A30138.
DR PDB; 1AJR; 20-AUG-97.
DR PDB; 1AJS; 20-AUG-97.
DR InterPro; IPR001511; Aminotran_1.

```

```

DR InterPro: IPR000796; Amino transferase_sub.
DR Pfam: PF00155; aminotran_1_2; 1.
DR PRINTS: PR00799; TRANSAMINASE.
DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
KW Transferase; Amino transferase; Pyridoxal phosphate; 3D-structure.
FT INIT_MET 0 0
FT ACT_SITE 40 40 IN THE PRESENCE OF SUBSTRATES, REACTS
FT WITH TETRAHYDROMETHANE WITH CONCOMITANT
FT INACTIVATION OF THE ENZYME.
FT BINDING 258 258 PYRIDOXAL PHOSPHATE.
FT CONFLICT 144 144 N -> D (IN REF. 2).
SQ SEQUENCE 412 AA; 46343 MW; 15FB716D2492E893 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 412;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 ARLEALK 206
Db 339 ARLEALK 345

RESULT 48
AATC_RAT
ID AATC_RAT STANDARD; PRT; 412 AA.
AC P13221; Q64570;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)
DE (Glutamate oxaloacetate transaminase-1).
GN GOT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89034271; PubMed=3182856;
RA Pave-Preux M., Ferry N., Bouguet J., Hanoune J., Barouki R.;
RT "Nucleotide sequence and glucocorticoid regulation of the mRNAs for
RT the isoenzymes of rat aspartate aminotransferase.";
RL J. Biol. Chem. 263:17459-17466(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89033975; PubMed=3053674;
RA Horio Y., Tanaka T., Taketoshi M., Nagashima F., Tanase S.,
RA Morino Y., Wada H.;
RT "Rat cytosolic aspartate aminotransferase: molecular cloning of cDNA
RT and expression in Escherichia coli.";
RL J. Biochem. 103:797-804(1988).
RN [3]
RP SEQUENCE OF 1-38 FROM N.A.
RX MEDLINE=90170950; PubMed=2307672;
RA Pave-Preux M., Aggerbeck M., Veyssier C., Bousquet-Lemerrier B.,
RA Hanoune J., Barouki R.;
RT "Hormonal discrimination among transcription start sites of aspartate
RT aminotransferase.";
RL J. Biol. Chem. 265:4444-4448(1990).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J04171; AAA40769.1; -
DR EMBL: D00252; BAA00183.1; -
DR EMBL: J05263; AAA40842.1; -
DR PIR: A32120; A32120.
DR PIR: J0439; J0439.
DR HSSP: P00503; 1AJS.
DR InterPro: IPR001511; Amino tran_1.
DR InterPro: IPR000796; Amino transferase_sub.
DR Pfam: PF00155; aminotran_1_2; 1.
DR PRINTS: PR00799; TRANSAMINASE.
DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
KW Transferase; Amino transferase; Pyridoxal phosphate.
FT INIT_MET 0 0
FT BINDING 258 258 PYRIDOXAL PHOSPHATE.
FT CONFLICT 54 54 T -> R (IN REF. 2).
FT CONFLICT 99 99 G -> R (IN REF. 2).
FT CONFLICT 108 108 A -> G (IN REF. 2).
FT CONFLICT 120 120 A -> G (IN REF. 2).
FT CONFLICT 124 124 N -> I (IN REF. 2).
FT CONFLICT 253 253 P -> A (IN REF. 2).
FT CONFLICT 365 365 L -> F (IN REF. 2).
SQ SEQUENCE 412 AA; 46197 MW; 5255D8F02B7334AB CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 412;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 ARLEALK 206
Db 339 ARLEALK 345

RESULT 49
AP04_MACFA
ID AP04_MACFA STANDARD; PRT; 429 AA.
AC P33621;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV).
GN APOA4.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Leukocyte;
RX MEDLINE=93192330; PubMed=8448212;
RA Osada J., Pocovi M., Nicolosi R.J., Schaefer E.J., Ordovas J.M.;
RT "Nucleotide sequences of the Macaca fascicularis apolipoprotein C-II
RT and A-IV genes.";
RL Biochim. Biophys. Acta 1172:335-339(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC HELICAL, AND MANY OF THESE HELICES ARE PREDICTED TO BE HIGHLY ALPHA-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECTININ:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.

```



```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68361; CAA48421.1; -.
DR PIR; S29565; S29565.
DR PIR; S30195; S30195.
DR HSSP; P02649; INFO.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 429 APOLIPOPROTEIN A-IV.
FT DOMAIN 33 330 13 X 22 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 33 54 1.
FT REPEAT 60 81 2.
FT REPEAT 82 103 3.
FT REPEAT 115 136 4.
FT REPEAT 137 158 5.
FT REPEAT 159 180 6.
FT REPEAT 181 202 7.
FT REPEAT 203 224 8.
FT REPEAT 225 246 9.
FT REPEAT 247 268 10.
FT REPEAT 269 286 11.
FT REPEAT 287 308 12.
FT REPEAT 309 330 13.
FT DOMAIN 372 420 GLU/GLN-RICH.
SQ SEQUENCE 429 AA; 49876 MW; 3DA58F551D0DB60C CRC64;

Query Match 2.6%; Score 7; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ELRQRLA 199
   |||||
Db 262 ELRQRLA 268

RESULT 50
YJUU_ECOLI STANDARD; PRT; 443 AA.
AC P39410;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yjjJ.
GN YJUU OR B4385.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [2]
RP SEQUENCE OF 1-49 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87259990; PubMed=3299264;
RA Larsen J.E.L., Albrechtsen B., Valentin-Hansen P.;
RT "Analysis of the terminator region after the deoCABD operon of
```

```
RT Escherichia coli K-12 using a new class of single copy number operon-
RT fusion vectors.";
RL Nucleic Acids Res. 15:5125-5140(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14003; AAA97281.1; -.
DR EMBL; AE000508; AAC77338.1; -.
DR EMBL; X05629; CAA29115.1; ALT_SEQ.
DR Ecogene; EG12342; yjjJ.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 443 AA; 49756 MW; 38B78B6921CEAC91 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ELRQRLA 199
   |||||
Db 17 ELRQRLA 23
```

Search completed: September 22, 2002, 12:27:09  
Job time: 229 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:02:49 ; Search time 51.49 Seconds  
(without alignments)  
80.644 Million cell updates/sec

Title: US-09-803-918a-2\_COPY\_25\_194  
Perfect score: 885  
Sequence: 1 DEPPQSPMDRVKDLATVYVD.....KARAHVDALRTHLAPYSDDEL 170

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	885	100.0	264	1	US-08-448-606-6 Sequence 6, Appl1
2	885	100.0	267	1	US-07-959-946-3 Sequence 3, Appl1
3	885	100.0	267	1	US-08-333-577-3 Sequence 3, Appl1
4	885	100.0	267	4	US-08-952-796-2 Sequence 2, Appl1
5	885	100.0	267	5	PCT-US92-08634-3 Sequence 3, Appl1
6	650	73.4	200	4	US-08-952-796-15 Sequence 15, Appl1
7	328	37.1	64	2	US-08-292-870-1 Sequence 1, Appl1
8	227	25.6	44	2	US-08-292-870-2 Sequence 2, Appl1
9	163	18.4	32	2	US-08-292-870-3 Sequence 3, Appl1
10	138	15.6	317	4	US-08-949-155-6 Sequence 6, Appl1
11	124	14.0	25	2	US-08-292-870-4 Sequence 4, Appl1
12	119.5	13.5	896	1	US-08-095-737-2 Sequence 2, Appl1
13	119.5	13.5	896	1	US-08-480-145-2 Sequence 2, Appl1
14	119.5	13.5	896	2	US-08-477-389-2 Sequence 2, Appl1
15	118.5	13.4	897	1	US-08-095-737-4 Sequence 4, Appl1
16	118.5	13.4	897	1	US-08-480-145-4 Sequence 4, Appl1
17	118.5	13.4	897	2	US-08-477-389-4 Sequence 4, Appl1
18	116.5	13.2	317	1	US-07-709-949-2 Sequence 2, Appl1
19	107.5	12.1	900	2	US-08-630-822A-62 Sequence 62, Appl1
20	107.5	12.1	900	2	US-09-005-069-62 Sequence 62, Appl1
21	106.5	12.0	107	1	US-08-182-175A-105 Sequence 105, Appl1
22	106.5	12.0	107	1	US-08-474-633A-92 Sequence 92, Appl1
23	106.5	12.0	107	5	PCT-US92-06412-105 Sequence 105, Appl1
24	106	12.0	515	2	US-08-705-660-46 Sequence 46, Appl1
25	106	12.0	515	3	US-08-989-045-46 Sequence 46, Appl1
26	101	11.4	424	2	US-08-951-148-9 Sequence 9, Appl1
27	101	11.4	424	2	US-09-165-234-9 Sequence 9, Appl1

28	101	11.4	424	3	US-09-274-570-9	Sequence 9, Appl1
29	100	11.3	110	1	US-07-849-389-7	Sequence 7, Appl1
30	100	11.3	220	2	US-08-726-306A-29	Sequence 29, Appl1
31	97.5	11.0	683	6	5210183-3	Patent No. 5210183
32	96	10.8	337	1	US-08-317-223-3	Sequence 3, Appl1
33	96	10.8	337	3	US-09-059-849A-3	Sequence 3, Appl1
34	96	10.8	337	3	US-09-213-632-3	Sequence 3, Appl1
35	96	10.8	337	5	PCT-US95-12675-3	Sequence 3, Appl1
36	96	10.8	3111	2	US-08-460-309-4	Sequence 4, Appl1
37	96	10.8	3111	2	US-08-125-077-4	Sequence 4, Appl1
38	95.5	10.8	428	1	US-07-882-790-4	Sequence 4, Appl1
39	95	10.7	337	1	US-08-445-135-4	Sequence 4, Appl1
40	95	10.7	1093	4	US-09-315-793-52	Sequence 52, Appl1
41	94	10.6	344	6	5210183-2	Patent No. 5210183
42	93	10.5	288	2	US-08-961-858-6	Sequence 6, Appl1
43	93	10.5	288	3	US-09-089-593-6	Sequence 6, Appl1
44	93	10.5	288	3	US-08-950-925-4	Sequence 4, Appl1
45	92	10.4	420	2	US-08-592-126-142	Sequence 142, Appl1

## ALIGNMENTS

RESULT 1  
US-08-448-606-6  
Sequence 6, Application US/08448606  
Patent No. 5721114  
GENERAL INFORMATION:  
APPLICANT: Abrahams n, Lars  
APPLICANT: Holmgren, Erik  
APPLICANT: Kalder n, Christina  
APPLICANT: Lake, Mats  
APPLICANT: Mikaelsson, sa  
APPLICANT: Sejlitz, Torsten  
TITLE OF INVENTION: Expression System For Producing  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pollock, Yande Sande & Priddy  
STREET: 1990 M Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,606  
FILING DATE: 25-AUG-1995  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE93/01061  
FILING DATE: 09-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9203753-0  
FILING DATE: 11-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Amernick, Burton A.  
REGISTRATION NUMBER: 24,852  
REFERENCE/DOCKET NUMBER: 0151/00121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)331-7111  
TELEFAX: (202) 293-6229  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-448-606-6

Query Match 100.0%; Score 885; DB 1; Length 264;  
Best Local Similarity 100.0%; Pred. No. 5e-72;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
|||||  
Db 22 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 81  
|||||

QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQEMELYRQKVE 120  
|||||  
Db 82 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQEMELYRQKVE 141  
|||||

QY 121 PLRAELQEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDEL 170  
|||||  
Db 142 PLRAELQEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDEL 191  
|||||

RESULT 2  
US-07-959-946-3  
; Sequence 3, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzlum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959,946  
; FILING DATE: 19921008  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-959-946-3

Query Match 100.0%; Score 885; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 5.1e-72;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
|||||

Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
|||||

QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQEMELYRQKVE 120  
|||||

Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQEMELYRQKVE 144  
|||||

QY 121 PLRAELQEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDEL 170  
|||||

Db 145 PLRAELQEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDEL 194  
|||||

RESULT 3  
US-08-333-577-3  
; Sequence 3, Application US/08333577  
; Patent No. 5786206  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzlum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; STREET: 180 No. 5786206th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,577  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER: SCRF 234.0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-333-577-3

Query Match 100.0%; Score 885; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 5.1e-72;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 60  
|||||

Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTSTFSKL 84  
|||||

QY 61 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQEMELYRQKVE 120  
|||||

Db 85 REQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQEMELYRQKVE 144  
|||||

QY 121 PLRAELQEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDEL 170  
|||||

Db 145 PLRAELQEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDEL 194  
|||||





APPLICANT: DUBERGER, Nicolas  
APPLICANT: FRUCHART, Jean-Charles  
APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPROTEIN A-1  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehlnert Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 200 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-15

Query Match 73.4%; Score 650; DB 4; Length 200;  
Best Local Similarity 99.2%; Pred. No. 4 le-51;  
Matches 126; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 LKLDNWDSTSTFSKRLQGLPVTOEFWDNLEKTEGLRQEMSKDLEEVKAKVQPYLDD 103  
Db 1 LKLDNWDSTSTFSKRLQGLPVTOEFWDNLEKTEGLRQEMSKDLEEVKAKVQPYLDD 60

QY 104 FQKKQOEEMELLYRQKVEPLRAELQEGARQKLTHELQEKLSPLGEEMDRARAHVDALRTYHL 163  
Db 61 FQKKQOEEMELLYRQKVEPLRAELQEGARQKLTHELQEKLSPLGEEMDRARAHVDALRTYHL 120

QY 164 APYSDEL 170  
Db 121 APYSDEL 127

RESULT 7  
US-08-292-870-1  
Sequence 1, Application US/08292870  
Patent No. 5814467  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS

TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESS: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCY/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-292-870-1

Query Match 37.1%; Score 328; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 7e-23;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 EMSKDEEVKAKVQPYLDDFQKKQOEEMELLYRQKVEPLRAELQEGARQKLTHELQEKLSPL 144  
Db 1 EMSKDEEVKAKVQPYLDDFQKKQOEEMELLYRQKVEPLRAELQEGARQKLTHELQEKLSPL 60

QY 145 GEEM 148  
Db 61 GEEM 64

RESULT 8  
US-08-292-870-2  
Sequence 2, Application US/08292870  
Patent No. 5814467  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSER: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-292-870-2

Query Match 25.6%; Score 227; DB 2; Length 44;  
Best Local Similarity 100.0%; Pred. No. 4.9e-14;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 RQMSKDLQVAKVQPYLDDFQKKQWQEMELYRQKVEPLRAEL 126  
Db 1 RQMSKDLQVAKVQPYLDDFQKKQWQEMELYRQKVEPLRAEL 44

RESULT 9  
US-08-292-870-3  
Sequence 3, Application US/08292870  
Patent No. 5814467  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSER: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11  
OTHER INFORMATION: /note= "Xaa can be either E (Glu)  
US-08-292-870-3

Query Match 18.4%; Score 163; DB 2; Length 32;  
Best Local Similarity 96.9%; Pred. No. 1.8e-08;  
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 KAKVQPYLDDFQKKQWQEMELYRQKVEPLRAE 125  
Db 1 KAKVQPYLDDFQKKQWQEMELYRQKVEPLRAE 32

RESULT 10  
US-08-949-155-6  
Sequence 6, Application US/08949155  
Patent No. 6271436  
GENERAL INFORMATION:  
APPLICANT: Piedrahita, Jorge A  
APPLICANT: Bazer, Fuller W  
TITLE OF INVENTION: Compositions and Methods for the  
TITLE OF INVENTION: Generation of Transgenic Animal Species  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE AND DURKEE  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: US  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/949,155  
FILING DATE: Concurrently Herewith

```

; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TAWK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-949-155-6

```

Query Match	15.6%;	Score 138;	DB 4;	Length 317;
Best Local Similarity	21.2%;	Pred. No. 5.5e-05;		
Matches 43;	Conservative 42;	Mismatches 78;	Indels 40;	Gaps 3;

[illegible]

RESULT 11  
US-08-292-870-4  
; Sequence 4, Application US/08292870  
; Patent No. 5814467  
; GENERAL INFORMATION:  
; APPLICANT: Curtiss, Linda K  
; APPLICANT: Banka, Carole L  
; APPLICANT: Bonnet, David J  
; APPLICANT: Smith, Richard S  
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/292,870  
; FILING DATE: 17-AUG-1994

```

; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-292-870-4

```

Query Match	14.08;	Score 124;	DB 2;	Length 25;
Best Local Similarity	96.08;	Pred. No. 4.1e-05;		
Matches 24;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

```
OY      87 SKDLEEVKAKVQPYLDDFOKKWQEE 1111  
         11111111111111 111111  
Db       1 SKDLEEVKAKVQPYLDDFOKKWQEE 25
```

```

RESULT 12
US-08-095-737-2
Sequence 2, Application US/08095737
Patent No. 5487979
GENERAL INFORMATION:
APPLICANT: DiFiore, Pier P
APPLICANT: Fazioli, Francesca
TITLE OF INVENTION: A Substrate for the Epidermal Growth
TITLE OF INVENTION: A Factor Receptor Kinase
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,737
FILING DATE: 19930722
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 896 amino acids
type: amino acid

```





```
Db      359 KEDTIKQRTSEVQDLQDEVQRENTNLQKIQAKQOVQELLDELDEQKAQLEQKVEVRKK 418
QY      108 WOHEMELYRQKVEPLRAEL--QEG-----ARQKLHEIQEKISPLGEEMDRARA 154
Db      419 CAEEAQL-----ISSLKAELTSQESQISTYEELIARKEELSRLOQETAEL-EESVESGKA 473
QY      155 HVDALRTHLAPYSDEL 170
Db      474 QLEPLOOHLQDSQOEI 489
```

## RESULT 15

```
US-08-095-737-4
; Sequence 4, Application US/08095737
; Patent No. 5487979
; GENERAL INFORMATION:
; APPLICANT: DiFiore, Pier P
; APPLICANT: Fazioli, Francesca
; TITLE OF INVENTION: A Substrate for the Epidermal Growth
; TITLE OF INVENTION: Factor Receptor Kinase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: United States of America
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,737
; FILING DATE: 19930722
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH060.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 897 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-095-737-4
```

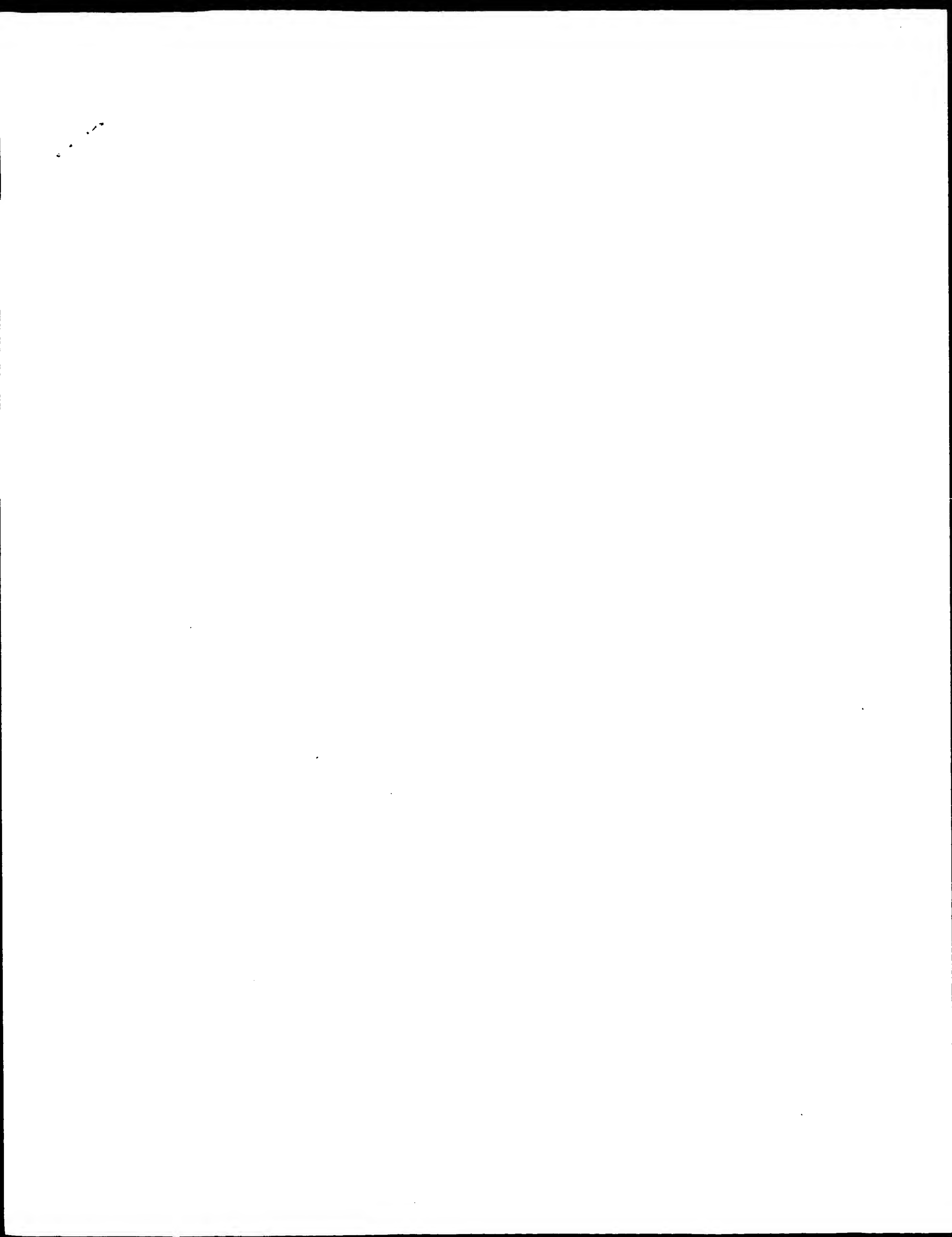
```
Query Match      13.48; Score 118.5; DB 1; Length 897;
Best Local Similarity 24.08; Pred. No. 0.011;
Matches 49; Conservative 43; Mismatches 65; Indels 47; Gaps 9;
```

```
QY      2 EPPQS-----PMDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLDNWDSTV 54
Db      298 DPPHSLTPEMIPPSDR-----SSLQKNI---TGSSPVADFSAIKELDTLNNEITVDLQREKN 350
QY      55 STFSKLRQQLGPV-----TQEFWDNLEKETEGRL-----QEMSKDLEEVKAKVQP 99
Db      351 NVEQDLKEKEDTVKQRTSEVQDLQDEVQRESINLQKIQAKQOVQELLGELDEQKAQLAE 410
QY      100 YLDDFOKKWOHEMELYRQKVEPLRAEL--QEG-----ARQKLHEIQEKISPLIGE 146
Db      411 QLGVEVRKKCAEEAQL-----ISSLKAELTSQESQISSYEELLKAREELSRLOQETAQL-E 465
QY      147 EMRDRARAHVDALRTHLAPYSDEL 170
Db      466 ESVESGKAQLPPLQOHLQESQOEI 489
```

Search completed: September 22, 2002, 12:02:50  
Job time: 250 sec

Sun Sep 22 12:10:28 2002

us-09-803-918a-2\_copy\_25\_194.ra





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 11:58:40 ; Search time 51.49 Seconds

(without alignments)  
126.658 Million cell updates/sec

Title: US-09-803-918A-2

Perfect score: 1362

Sequence: 1 MKAVALTLAVLFLTGSGARH.....SEKVSFLSALFEYTKRLNTQ 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	267	1	US-07-959-946-3 Sequence 3, Appli
2	1362	100.0	267	1	US-08-333-577-3 Sequence 3, Appli
3	1362	100.0	267	4	US-08-952-796-2 Sequence 2, Appli
4	1362	100.0	267	5	PCT-US92-08634-3 Sequence 3, Appli
5	1240.5	91.1	264	1	US-08-448-606-6 Sequence 6, Appli
6	1006	73.9	200	4	US-08-952-796-15 Sequence 15, Appli
7	328	24.1	64	2	US-08-292-870-1 Sequence 1, Appli
8	227	16.7	44	2	US-08-949-155-6 Sequence 6, Appli
9	190	14.0	317	4	US-07-709-949-2 Sequence 2, Appli
10	171.5	12.6	317	1	US-08-292-870-3 Sequence 3, Appli
11	163	12.0	32	2	US-08-630-822A-62 Sequence 62, Appli
12	129	9.5	900	2	US-09-005-069-62 Sequence 29, Appli
13	129	9.5	900	2	US-08-726-306A-29 Sequence 4, Appli
14	126	9.3	220	2	US-08-292-870-4 Patent No. 5210183
15	124	9.1	25	6	US-08-095-737-2 Sequence 2, Appli
16	123.5	9.0	896	1	US-08-480-145-2 Sequence 2, Appli
17	122.5	9.0	896	2	US-08-477-389-2 Sequence 11, Appli
18	122.5	9.0	896	2	US-08-095-737-4 Sequence 4, Appli
19	122.5	9.0	897	1	US-08-477-389-4 Sequence 4, Appli
20	122	9.0	1588	5	PCT-US93-07261-16 Sequence 16, Appli
21	122	9.0	1663	5	PCT-US93-07261-16 Sequence 16, Appli
22	121.5	8.9	897	1	US-08-477-389-4 Sequence 4, Appli
23	121.5	8.9	897	1	US-08-477-389-4 Sequence 4, Appli
24	121.5	8.9	897	1	US-08-477-389-4 Sequence 4, Appli
25	121	8.9	1184	4	US-09-541-782-2 Sequence 2, Appli
26	119	8.7	337	1	US-08-445-135-4 Sequence 4, Appli
27	118	8.7	337	1	US-08-317-223-3 Sequence 3, Appli

28	118	8.7	337	3	US-09-059-849A-3	Sequence 3, Appli
29	118	8.7	337	3	US-09-213-632-3	Sequence 3, Appli
30	118	8.7	337	5	PCT-US95-12675-3	Sequence 3, Appli
31	118	8.7	3111	2	US-08-460-309-4	Sequence 4, Appli
32	118	8.7	3111	2	US-08-125-077-4	Sequence 4, Appli
33	117.5	8.6	428	1	US-07-882-790-4	Sequence 4, Appli
34	117.5	8.6	515	2	US-08-705-660-46	Sequence 46, Appli
35	117.5	8.6	515	3	US-08-989-045-46	Sequence 46, Appli
36	117	8.6	1786	4	US-08-973-462-8	Sequence 8, Appli
37	117	8.6	3248	5	US-08-353-700-1	Sequence 1, Appli
38	117	8.6	3248	5	PCT-US95-16216-1	Sequence 52, Appli
39	116	8.5	1093	4	US-09-315-793-52	Sequence 4, Appli
40	116	8.5	2101	1	US-08-466-390-4	Sequence 4, Appli
41	116	8.5	2101	1	US-08-470-950-4	Sequence 4, Appli
42	116	8.5	2101	1	US-08-467-781-4	Sequence 4, Appli
43	116	8.5	2101	1	US-08-195-487-4	Sequence 4, Appli
44	116	8.5	2101	2	US-08-483-924-4	Sequence 4, Appli
45	116	8.5	2101	4	US-09-452-294-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-07-959-946-3  
; Sequence 3, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959, 946  
; FILING DATE: 19921008  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901, 706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamsen, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-946-3  
Query Match 100.0%; Score 1362; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 3.1e-109;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVCLKDSGRDYVSQFECS 60  
DB 1 MKAAYLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVCLKDSGRDYVSQFECS 60  
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGAROKLHELQEKLSPLGGEEMDRARAHV 180  
DB 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGAROKLHELQEKLSPLGGEEMDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSLSEKAKPALLEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSLSEKAKPALLEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267  
DB 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267

## RESULT 2

US-08-333-577-3  
; Sequence 3, Application US/08333577  
; Patent No. 5786206

## ; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzlum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5786206th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,577  
; FILING DATE:

## ; CLASSIFICATION: 530

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Gamson, Edward P.

; REGISTRATION NUMBER: 29,381

; REFERENCE/DOCKET NUMBER: SCRF 234.0

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312)616-5400

; TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-333-577-3

Query Match 100.0%; Score 1362; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.1e-109;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVCLKDSGRDYVSQFECS 60  
DB 1 MKAAYLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVCLKDSGRDYVSQFECS 60  
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGAROKLHELQEKLSPLGGEEMDRARAHV 180  
DB 121 VQPYLDDFQKKWQEMELLYRQKVEPLRAELQEGAROKLHELQEKLSPLGGEEMDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSLSEKAKPALLEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELYHAKATEHLSLSEKAKPALLEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267  
DB 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267

## RESULT 3

US-08-952-796-2  
; Sequence 2, Application US/08952796  
; Patent No. 6258596

## ; GENERAL INFORMATION:

; APPLICANT: BENOTT, Patrick  
; APPLICANT: BRUCKERT, Eric  
; APPLICANT: DENEFFLE, Patrice  
; APPLICANT: DUBERGER, Nicolas  
; APPLICANT: FRUCHART, Jean-Charles  
; APPLICANT: LUC, Gerald  
; APPLICANT: TURPIN, Gerard  
; APPLICANT: ASSMANN, Gerd  
; APPLICANT: FUNKE, Harald  
; TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPROTEIN A-I  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, Mailstop 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/952,796  
; FILING DATE:

## ; CLASSIFICATION: 424

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 95/06061

; FILING DATE: 22-MAY-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO FR96/00747

; FILING DATE: 20-MAY-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fehlner Esq., Paul F.

; REGISTRATION NUMBER: 35,135

; REFERENCE/DOCKET NUMBER: ST95031-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (610) 454-3839

; TELEFAX: (610) 454-3808

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-952-796-2

Query Match 100.0%; Score 1362; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.1e-109;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEES 60  
DB 1 MKAAYLTAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEES 60  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
DB 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267  
DB 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267

## RESULT 4

PCT-US92-08634-3

; Sequence 3, Application PC/RUS9208634

; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.

; APPLICANT: Koduri, Raju

; APPLICANT: Young, Stephen G.

; APPLICANT: Witzum, Joseph L.

; APPLICANT: Curtiss, Linda K.

; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a

; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &amp;

; ADDRESSEE: Milnamow, Ltd.

; STREET: 180 North Stetson, Suite 4700

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/08634

; FILING DATE: 19921009

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/901,706

; FILING DATE: 18-JUN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Gamsen, Edward P.

; REGISTRATION NUMBER: 29,381

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312)616-5400

; TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US92-08634-3

Query Match 100.0%; Score 1362; DB 5; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.1e-109;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEES 60  
DB 1 MKAAYLTAVLFLTGSQARHFHQODEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEES 60  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
DB 121 VQPYLDDFQKKWQEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267  
DB 241 GLLPVLESFKVSFLSALAEYTKKLNTQ 267

## RESULT 5

US-08-448-606-6

; Sequence 6, Application US/08448606

; Patent No. 5721114

; GENERAL INFORMATION:

; APPLICANT: Abrahams n, Lars

; APPLICANT: Holmgren, Erik

; APPLICANT: Kalder n, Christina

; APPLICANT: Lake, Mats

; APPLICANT: Mikaelsson, sa

; APPLICANT: Sejlitz, Torsten

; TITLE OF INVENTION: Expression System For Producing

; TITLE OF INVENTION: Apolipoprotein AI-M

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pollock, Vande Sande &amp; Priddy

; STREET: 1990 M Street, N.W., Suite 800

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448,606

; FILING DATE: 25-AUG-1995

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/SE93/01061

; FILING DATE: 09-DEC-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: SE 9203753-0

; FILING DATE: 11-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Amernick, Burton A.

; REGISTRATION NUMBER: 24,852

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)331-7111

; TELEFAX: (202)293-6229

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:





ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-292-870-1

Query Match 24.1%; Score 328; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 1.6e-21;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 EMSKDLSEVKKAKVQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLTLEQKLSPL 168  
|||||  
Db 1 EMSKDLSEVKKAKVQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLTLEQKLSPL 60

QY 169 GEEM 172  
|||||  
Db 61 GEEM 64

## RESULT 8

US-08-292-870-2  
Sequence 2, Application US/08292870  
Patent No. 5814467

## GENERAL INFORMATION:

APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994

## CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991

## ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-292-870-2

Query Match 16.7%; Score 227; DB 2; Length 44;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 RQMSKDLSEVKKAKVQPYLDDFQKKWQEMELYRQKVEPLRAEL 150  
|||||  
Db 1 RQMSKDLSEVKKAKVQPYLDDFQKKWQEMELYRQKVEPLRAEL 44

## RESULT 9

US-08-949-155-6  
Sequence 6, Application US/08949155  
Patent No. 6271436

## GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A  
APPLICANT: Bazer, Fuller W  
TITLE OF INVENTION: Compositions and Methods for the  
Generation of Transgenic Animal Species  
TITLE OF INVENTION: Generation of Transgenic Animal Species  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE AND DURKEE  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: US  
ZIP: 77210-4433

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/949,155  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/027,338  
FILING DATE: 11-OCT-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,094  
FILING DATE: 09-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hibler, David W.  
REGISTRATION NUMBER: 41,071  
REFERENCE/DOCKET NUMBER: TAMK:177  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679

## INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 317 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-949-155-6

Query Match 14.0%; Score 190; DB 4; Length 317;  
Best Local Similarity 21.9%; Pred. No. 9e-09;  
Matches 68; Conservative 60; Mismatches 112; Indels 70; Gaps 8;





SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11  
OTHER INFORMATION: /note= "Xaa can be either E (Glu)  
OTHER INFORMATION: or F (Phe)"  
US-08-292-870-3

Query Match 12.0%; Score 163; DB 2; Length 32;  
Best Local Similarity 96.9%; Pred. No. 8.7e-08;  
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 118 KAKVQPYLDYDQKKWQEMELLYRQKVEPLRAE 149  
DB 1 KAKVQPYLDYDQKKWQEMELLYRQKVEPLRAE 32

## RESULT 12

US-08-630-822A-62  
Sequence 62, Application US/08630822A  
Patent No. 5840695

## GENERAL INFORMATION:

APPLICANT: FRANK, GLENN R.  
APPLICANT: HUNTER, SHIRLEY WU  
APPLICANT: WALLENFELS, LYNDY  
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,822A  
FILING DATE: 11-APR-1996  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: CONNELL, GARY J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-17-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 900 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-630-822A-62

Query Match 9.5%; Score 129; DB 2; Length 900;  
Best Local Similarity 24.3%; Pred. No. 0.0061;  
Matches 57; Conservative 46; Mismatches 80; Indels 52; Gaps 9;

QY 51 RDYVSQFEGSALGKQLNLKLDN-WDSVTSTFSKLRQLGPTQ----- 93

DB 388 QNLLQEKGNSENKLTNTQLLSNKLDELGQRECELRNQAAGDYKEKELTKFKLSCKELQKRA 447

QY 94 EFWDNLEKETEGLRQEMSKDLEEVKAK-----VQPYLDDFQKKWQEM 136  
DB 448 EFENELRRKTESLIVETKKRLDEEQNKRTREMNNNQHNKINMLEKQINDLQEKLGEL 507  
QY 137 ELYRQKVEPLRAEL---QEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLA----- 188  
DB 508 E-HNOKLKKQAVELRYVAQATEQLNNELQETMGL-QTORDALQOEVAASLGKLSQERSS 565  
QY 189 -PYSDELRLQRLAELK-----ENGARLAEYHAKATEHLSTSEKAPALE 236  
DB 566 RSQASDMQIELEAKLQALHLELHVNRNCEPKVTQDNRLRLERISTL-EKECASLE 619

RESULT 13  
US-09-005-069-62  
Sequence 62, Application US/09005069  
Patent No. 5932470

## GENERAL INFORMATION:

APPLICANT: FRANK, GLENN R.  
APPLICANT: HUNTER, SHIRLEY WU  
APPLICANT: WALLENFELS, LYNDY  
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,069  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/630,822  
FILING DATE: 11-APR-1996  
ATTORNEY/AGENT INFORMATION:

NAME: CONNELL, GARY J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-17-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 900 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-005-069-62

Query Match 9.5%; Score 129; DB 2; Length 900;  
Best Local Similarity 24.3%; Pred. No. 0.0061;  
Matches 57; Conservative 46; Mismatches 80; Indels 52; Gaps 9;

QY 51 RDYVSQFEGSALGKQLNLKLDN-WDSVTSTFSKLRQLGPTQ----- 93

DB 388 QNLLQEKGNSENKLTNTQLLSNKLDELGQRECELRNQAAGDYKEKELTKFKLSCKELQKRA 447

QY 94 EFWDNLEKETEGLRQEMSKDLEEVKAK-----VQPYLDDFQKKWQEM 136

DB 448 EFENELRRKTESLIVETKKRLDEEQNKRTREMNNNQHNKINMLEKQINDLQEKLGEL 507

QY 137 ELYRQKVEPLRAEL---QEGARQKLHELQEKLSPLGEMRDRARAHVDALRTHLA----- 188

Db 508 E-HNOKLKQAVEELRYAQSATEQLNLEQETMQL-QTORDALQOEYASLQKLSQERS 565  
QY 189 -PYSDELRLRLARLEALK-----ENGARLAHYHAKATEHLSTLSEKAKPALE 236  
Db 566 RSQASDMQIELEAKLQALHIELEHVNRNCEKVTQDNRLRLERISTL-EKECASLE 619

## RESULT 14

US-08-726-306A-29

; Sequence 29, Application US/08726306A

; Patent No. 5958684

; GENERAL INFORMATION:

; APPLICANT: van Leeuwen, Frederik Willem

; APPLICANT: Burbach, Johannes Peter Henri

; APPLICANT: Grosveld, Franklin G.

; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS

; NUMBER OF SEQUENCES: 189

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner &amp; Witcoff, Ltd.

; STREET: 1 Financial Center

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Wordperfect 6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/726,306A

; FILING DATE: 02-Oct-1996

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: GB 95/20080.4

; FILING DATE: 02-Oct-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 60/009,832

; FILING DATE: 01-Jan-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams, Ph.D., Kathleen M.

; REGISTRATION NUMBER: 34,380

; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 345-9100

; TELEFAX: (617) 345-9111

; INFORMATION FOR SEQ ID NO: 29:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 220 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-726-306A-29

Query Match 9.3%; Score 126; DB 2; Length 220;

Best Local Similarity 22.2%; Pred. No. 0.0017;

Matches 51; Conservative 34; Mismatches 95; Indels 50; Gaps 3;

QY 1 MKAVALTLAVLFTGSOARHFWQDEPPQSP-----WDRVKDLAT 40  
Db 21 MKVLMAALLVTLFAGCQAKVEQAVETEPEPELRQQTWQSGQRMWELALGRFWDYLRWQT 80  
QY 41 VYVDVLKDSGRDYSQFEGSALGKQNLKLDNWDVSTSTPSKLRQQLGPTQEFWDNLE 100  
Db 81 LSEQVQELLSQVYQ-----ELRALMDETMKELKAYNSELEQQLTPVAETRARLS 132  
QY 101 KETEGLRQENSKDLSEVKAKVQPYLDLDFQKKQWQEMELYRQKVEPLRAELQEGARQKLHE 160  
Db 133 KELQTAQARLQADMEDVCGRLVQY-----RGEVQAMLGQSTEE 170  
QY 161 LOEKLSPGEMRDRARAHVDALRTHLAPYSDELRLRLARLEALKENG 210

Db 171 LRVRLASHLRKRLRLLRDPDDLQKRLAVYQAGAREGSAIRERLG 220

## RESULT 15

US-08-292-870-4

; Sequence 4, Application US/08292870

; Patent No. 5814467

; GENERAL INFORMATION:

; APPLICANT: Curtiss, Linda K

; APPLICANT: Banka, Carole L

; APPLICANT: Bonnet, David J

; APPLICANT: Smith, Richard S

; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS

; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Scripps Research Institute, Office of

; ADDRESSEE: Patent Counsel

; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8

; CITY: La Jolla

; STATE: California

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/292,870

; FILING DATE: 17-AUG-1994

; CLASSIFICATION: 436

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 07/534,761

; FILING DATE: 07-JUN-1990

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 07/711,333

; FILING DATE: 06-JUN-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

; REGISTRATION NUMBER: 34,163

; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-554-2937

; TELEFAX: 619-554-6312

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

; US-08-292-870-4

Query Match 9.1%; Score 124; DB 2; Length 25;

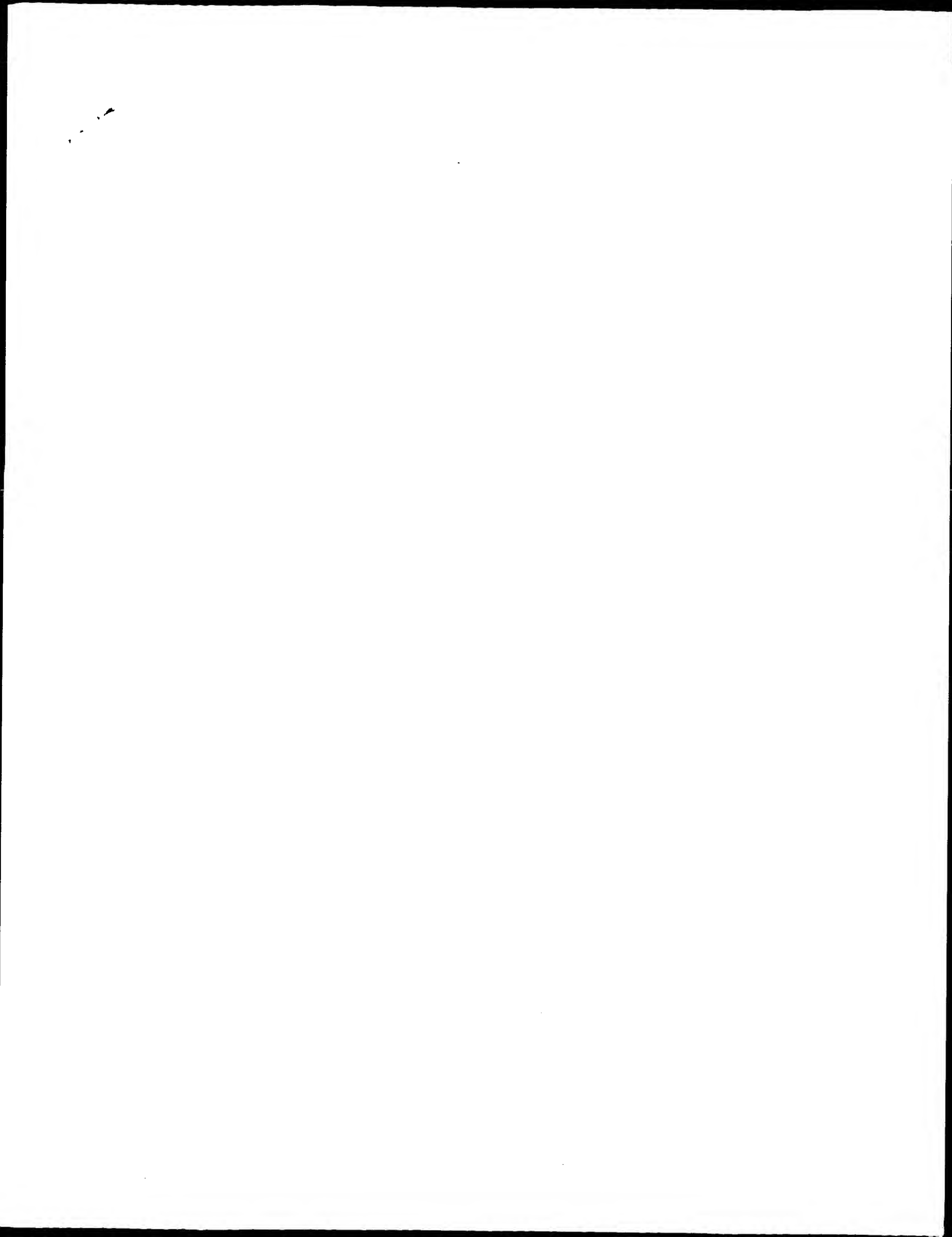
Best Local Similarity 96.0%; Pred. No. 0.00014;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 111 SKDLEEVAKVQPYLDLDFQKKWQEE 135  
Db 1 SKDLEEVAKVQPYLDLDFQKKWQEE 25

Search completed: September 22, 2002, 12:02:49  
Job time: 249 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 12:16:40 ; Search time 13.02 Seconds  
(without alignments)  
500.894 Million cell updates/sec

Title: US-09-803-918A-2  
Perfect score: 267  
Sequence: 1 MKAAVLTAVLFLTGSQARH.....SFKVSFLSALEETKRLNTQ 267

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	100.0	267	1	US-07-959-946-3 Sequence 3, Appli
2	267	100.0	267	1	US-08-333-577-3 Sequence 3, Appli
3	267	100.0	267	4	US-08-952-796-2 Sequence 2, Appli
4	267	100.0	267	5	PCT-US92-08634-3 Sequence 3, Appli
5	172	64.4	264	1	US-08-448-606-6 Sequence 6, Appli
6	107	40.1	200	4	US-08-952-796-15 Sequence 15, Appli
7	64	24.0	64	2	US-08-292-870-1 Sequence 1, Appli
8	44	16.5	44	2	US-08-292-870-2 Sequence 2, Appli
9	21	7.9	32	2	US-08-292-870-3 Sequence 3, Appli
10	17	6.4	25	2	US-08-292-870-4 Sequence 4, Appli
11	16	6.0	16	1	US-07-959-946-5 Sequence 5, Appli
12	16	6.0	16	1	US-08-333-577-5 Sequence 5, Appli
13	16	6.0	16	5	PCT-US92-08634-5 Sequence 5, Appli
14	9	3.4	13	5	PCT-US94-01234-47 Sequence 47, Appli
15	8	3.0	8	2	US-08-630-645-6 Sequence 6, Appli
16	8	3.0	8	5	PCT-US96-10220-6 Sequence 6, Appli
17	8	3.0	317	4	US-08-949-155-6 Sequence 6, Appli
18	7	2.6	7	4	US-08-952-796-11 Sequence 11, Appli
19	7	2.6	13	1	US-08-448-606-2 Sequence 2, Appli
20	7	2.6	70	3	US-09-101-146-4 Sequence 4, Appli
21	7	2.6	229	4	US-09-188-930-315 Sequence 315, App
22	7	2.6	233	4	US-09-188-930-139 Sequence 139, App
23	7	2.6	315	4	US-09-180-827-6 Sequence 2, Appli
24	7	2.6	343	1	US-08-336-891-2 Sequence 2, Appli
25	7	2.6	343	5	PCT-US95-13795-4 Sequence 4, Appli
26	7	2.6	482	4	US-08-637-823B-32 Sequence 32, Appli
27	7	2.6	483	4	US-08-637-823B-4 Sequence 4, Appli

#### ALIGNMENTS

28	7	2.6	484	4	US-08-637-823B-2	Sequence 2, Appli
29	7	2.6	547	3	US-08-903-139B-7	Sequence 7, Appli
30	7	2.6	548	2	US-08-676-279-50	Sequence 50, Appli
31	7	2.6	548	3	US-08-903-139B-8	Sequence 8, Appli
32	7	2.6	548	3	US-08-903-139B-9	Sequence 9, Appli
33	7	2.6	548	3	US-08-903-139B-28	Sequence 28, Appli
34	7	2.6	548	4	US-08-637-823B-25	Sequence 25, Appli
35	7	2.6	549	2	US-08-676-279-57	Sequence 57, Appli
36	7	2.6	550	4	US-08-637-823B-30	Sequence 30, Appli
37	7	2.6	568	4	US-08-637-823B-27	Sequence 27, Appli
38	7	2.6	584	4	US-08-637-823B-28	Sequence 28, Appli
39	7	2.6	629	4	US-09-300-909-19	Sequence 19, Appli
40	7	2.6	882	4	US-09-413-814-78	Sequence 78, Appli
41	7	2.6	2254	2	US-08-677-010-3	Sequence 3, Appli
42	7	2.6	2254	2	US-08-790-519-3	Sequence 3, Appli
43	7	2.6	2509	1	US-08-469-005A-10	Sequence 10, Appli
44	7	2.6	2511	4	US-09-261-907-2	Sequence 2, Appli
45	7	2.6	5087	4	US-09-144-085-1	Sequence 1, Appli
46	6	2.2	7	1	US-08-704-170-43	Sequence 43, Appli
47	6	2.2	7	2	US-08-474-741-1	Sequence 1, Appli
48	6	2.2	7	4	US-09-187-859-1184	Sequence 1184, Ap
49	6	2.2	7	5	PCT-US94-02631-43	Sequence 43, Appli
50	6	2.2	8	4	US-09-187-859-1147	Sequence 1147, Ap
51	6	2.2	8	4	US-09-187-859-1189	Sequence 1189, Ap
52	6	2.2	10	4	US-09-187-859-1150	Sequence 1150, Ap
53	6	2.2	10	4	US-09-187-859-1192	Sequence 1192, Ap
54	6	2.2	12	1	US-08-117-362-28	Sequence 28, Appli
55	6	2.2	12	1	US-08-486-924-28	Sequence 28, Appli
56	6	2.2	15	1	US-08-218-025A-144	Sequence 144, Appli
57	6	2.2	17	4	US-08-954-395A-4	Sequence 4, Appli
58	6	2.2	18	3	US-08-940-095-208	Sequence 208, App
59	6	2.2	18	3	US-08-940-093-208	Sequence 208, App
60	6	2.2	18	3	US-08-940-096-208	Sequence 208, App
61	6	2.2	18	4	US-09-465-719-208	Sequence 208, App
62	6	2.2	18	4	US-09-453-605-208	Sequence 208, App
63	6	2.2	20	1	US-08-484-635-140	Sequence 140, App
64	6	2.2	20	2	US-08-484-631-140	Sequence 140, App
65	6	2.2	20	2	US-08-827-570-140	Sequence 140, App
66	6	2.2	22	1	US-08-484-635-134	Sequence 134, App
67	6	2.2	22	1	US-08-484-635-144	Sequence 144, App
68	6	2.2	22	2	US-08-484-631-134	Sequence 134, App
69	6	2.2	22	2	US-08-484-631-144	Sequence 144, App
70	6	2.2	22	2	US-08-827-570-134	Sequence 134, App
71	6	2.2	22	2	US-08-827-570-144	Sequence 144, App
72	6	2.2	22	3	US-08-940-095-102	Sequence 102, App
73	6	2.2	22	3	US-08-940-095-109	Sequence 109, App
74	6	2.2	22	3	US-08-940-095-190	Sequence 190, App
75	6	2.2	22	3	US-08-940-093-102	Sequence 102, App

RESULT 1  
US-07-959-946-3  
; Sequence 3, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witzlum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA

ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/959,946  
FILING DATE: 19921008  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-959-946-3

Query Match 100.0%; Score 267; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.7e-250;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVLYKDSGRDYVSQFEGS 60  
DB 1 MKAAVLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVLYKDSGRDYVSQFEGS 60  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDEQKKWQEMELRYRQKVEPLRAELQEGAROKLHLEQKLSPLGEMRDRARAHV 180  
DB 121 VQPYLDDEQKKWQEMELRYRQKVEPLRAELQEGAROKLHLEQKLSPLGEMRDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHISTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHISTLSEKAKPALEDLRQ 240  
QY 241 GLPVLESFVKVSLALEEYTKKLNQ 267  
DB 241 GLPVLESFVKVSLALEEYTKKLNQ 267

RESULT 2  
US-08-333-577-3  
Sequence 3, Application US/08333577  
Patent No. 5786206

GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Witzum, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 No. 5786206th stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois

COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,577  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: SCRF 234.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-333-577-3

Query Match 100.0%; Score 267; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.7e-250;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVLYKDSGRDYVSQFEGS 60  
DB 1 MKAAVLTAVLFLTGSQARHFWQODEPPQSPWDRVKDLATVYVDVLYKDSGRDYVSQFEGS 60  
QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGKQLNLKLDNWDVSTSTFSKLRQGLGPTVQEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDEQKKWQEMELRYRQKVEPLRAELQEGAROKLHLEQKLSPLGEMRDRARAHV 180  
DB 121 VQPYLDDEQKKWQEMELRYRQKVEPLRAELQEGAROKLHLEQKLSPLGEMRDRARAHV 180  
QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHISTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAEYHAKATEHISTLSEKAKPALEDLRQ 240  
QY 241 GLPVLESFVKVSLALEEYTKKLNQ 267  
DB 241 GLPVLESFVKVSLALEEYTKKLNQ 267

RESULT 3  
US-08-952-796-2  
Sequence 2, Application US/08952796  
Patent No. 6258596

GENERAL INFORMATION:  
APPLICANT: BENOIT, Patrick  
APPLICANT: BRUCKERT, Eric  
APPLICANT: DENEFFLE, Patrice  
APPLICANT: DUBERGER, Nicolas  
APPLICANT: FRUCHART, Jean-Charles  
APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-1  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA

COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehlner Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-2

Query Match 100.0%; Score 267; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.7e-250;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSAARHFHQDEPPQSPWDRVKDLATVYVDVLDKDSGRDYVSQFEES 60  
DB 1 MKAAYLTAVLFLTGSAARHFHQDEPPQSPWDRVKDLATVYVDVLDKDSGRDYVSQFEES 60  
QY 61 ALGQNLKLLDNWDSVSTFSKLRQGLPVTOEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGQNLKLLDNWDSVSTFSKLRQGLPVTOEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFOKKWQEMELRYKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
DB 121 VQPYLDDFOKKWQEMELRYKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
QY 181 DALRTHLAPYSDELQRRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELQRRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALEEYTKKLTQ 267  
DB 241 GLLPVLESFKVSFLSALEEYTKKLTQ 267

RESULT 4  
PCT-US92-08634-3  
Sequence 3, Application PC/TUS9208634  
GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Wiltum, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.

STREET: 180 North Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08634  
FILING DATE: 19921009  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamsen, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-08634-3

Query Match 100.0%; Score 267; DB 5; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.7e-250;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAYLTAVLFLTGSAARHFHQDEPPQSPWDRVKDLATVYVDVLDKDSGRDYVSQFEES 60  
DB 1 MKAAYLTAVLFLTGSAARHFHQDEPPQSPWDRVKDLATVYVDVLDKDSGRDYVSQFEES 60  
QY 61 ALGQNLKLLDNWDSVSTFSKLRQGLPVTOEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
DB 61 ALGQNLKLLDNWDSVSTFSKLRQGLPVTOEFWDNLEKETEGLRQEMSKDLEEVKAK 120  
QY 121 VQPYLDDFOKKWQEMELRYKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
DB 121 VQPYLDDFOKKWQEMELRYKVEPLRAELQEGAROKLHELQEKLSPLGEMRDRARAHV 180  
QY 181 DALRTHLAPYSDELQRRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
DB 181 DALRTHLAPYSDELQRRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240  
QY 241 GLLPVLESFKVSFLSALEEYTKKLTQ 267  
DB 241 GLLPVLESFKVSFLSALEEYTKKLTQ 267

RESULT 5  
US-08-448-606-6  
Sequence 6, Application US/08448606  
Patent No. 5721114  
GENERAL INFORMATION:  
APPLICANT: Abrahams n, Lars  
APPLICANT: Holmgren, Erik  
APPLICANT: Kaldner n, Christina  
APPLICANT: Lake, Mats  
APPLICANT: Mikaelsson, sa  
APPLICANT: Sejlitz, Torsten  
TITLE OF INVENTION: Expression System For Producing  
TITLE OF INVENTION: Apolipoprotein AI-M  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:

```

1 ADDRESSSEE: Pollock, Vande Sande & Priddy
2 STREET: 1990 M Street, N.W., Suite 800
3 CITY: Washington
4 STATE: D.C.
5 COUNTRY: U.S.
6 ZIP: 20036
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Floppy disk
10 COMPUTER: IBM PC compatible
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12 SOFTWARE: PatentIn Release #1.0, Version #1.30
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/448,606
16 FILING DATE: 25-AUG-1995
17
18 CLASSIFICATION: 436
19
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: PCT/SE93/01061
22 FILING DATE: 09-DEC-1993
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: SE 9203753-0
26 FILING DATE: 11-DEC-1992
27
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Amernick, Burton A.
30 REGISTRATION NUMBER: 24,852
31 REFERENCE/DOCKET NUMBER: 0151/00121
32
33 TELECOMMUNICATION INFORMATION:
34 TELEPHONE: (202)331-7111
35 TELEFAX: (202) 293-6229
36
37 INFORMATION FOR SEQ ID NO: 6:
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 264 amino acids
40 TYPE: amino acid
41
42 TOPOLOGY: linear
43
44 MOLECULE TYPE: protein
45
46 US-08-448-606-6

```

Query Match	64.4%;	Score 172;	DB 1;	Length 264;
Best Local Similarity	100.0%;	Pred. No. 2.2e-158;		
Matches 172;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	25	DEPPOSFMDRVKDLATVYVDVLKDSGRDVSQFEGSALGKQLNLKLDNWDSVTSTFSKL	84	
Db	22	DEPPQSPMDRVKDLATVYVDVLKDSGRDVSQFEGSALGKQLNLKLDNWDSVTSTFSKL	81	
QY	85	REQLGPVTOEFWMDLLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQGEEMELYRQKVE	144	
Db	82	REQLGPVTOEFWMDLLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKQGEEMELYRQKVE	141	
QY	145	PLRAELQEGAROKLHELQEKISPLGGEEMDRARAHVDALRTHLAPYSDELRO	196	
Db	142	PLRAELQEGAROKLHELQEKISPLGGEEMDRARAHVDALRTHLAPYSDELRO	193	

RESULT 6  
 US-08-952-796-15  
 ; Sequence 15, Application US/08952796  
 ; Patent No. 6258596  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BENOIT, Patrick  
 ; APPLICANT: BROCKERT, Eric  
 ; APPLICANT: DENEFE, Patrice  
 ; APPLICANT: DUBERGER, Nicolas  
 ; APPLICANT: FRUCHART, Jean-Charles  
 ; APPLICANT: LUC, Gerald  
 ; APPLICANT: TURPIN, Gerrard  
 ; APPLICANT: ASSMANN, Gerd  
 ; APPLICANT: FUNKE, Harald  
 ; TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 ; STREET: 500 Arcola Road, Mailstop 3C43

```

: CITY: Collegeville
: STATE: PA
: COUNTRY: USA
: ZIP: 19426
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/952,796
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 95/06061
: FILING DATE: 22-MAY-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO FR96/00747
: FILING DATE: 20-MAY-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Fehlner Esq., Paul F.
: REGISTRATION NUMBER: 35,135
: REFERENCE/DOCKET NUMBER: ST95031-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (610) 454-3839
: TELEFAX: (610) 454-3808
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 200 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-952-796-15

```

	Query Match	40.1%;	Score 107;	DB 4;	length 200;	
	Best Local Similarity	100.0%;	Pred. No. 1.1e-95;			
	Matches 107;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	68 LKLLDNWDSVTSTFSKLRQLGVPVTFEFWDNLEKETEGTLRQEMSKDLEEVKAKVQPYLDD	127				
Dd	1 LKLLDNWDSVTSTFSKLRQEGPVTQFEFWDNLEKETEGTLRQEMSKDLEEVKAKVQPYLDD	60				
QY	128 FQKKWOEEMELYRQKVEPLRAELQEGAROKKLHLEQLKSLPIGEEMRD	174				
Dd	61 FQKKWOEEMELYRQKVEPLRAELQEGAROKKLHLEQLKSLPIGEEMRD	107				

RESULT 7  
 US-08-292-870-1  
 ; Sequence 1, Application US/08292870  
 ; Patent No. 5814467  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Curtiss, Linda K  
 ; APPLICANT: Banka, Carole L  
 ; APPLICANT: Bonnet, David J  
 ; APPLICANT: Smith, Richard S  
 ; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
 ; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
 ; TITLE OF INVENTION: METHODS  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: The Scripps Research Institute, Office of  
 ; ADDRESSEE: Patent Counsel  
 ; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
 ; CITY: La Jolla  
 ; STATE: California  
 ; COUNTRY: US  
 ; ZIP: 92037  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS



```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,870
FILING DATE: 17-AUG-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/534,761
FILING DATE: 07-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/711,333
FILING DATE: 06-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 91/04038
FILING DATE: 07-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-292-870-1
```

```
Query Match      24.0%; Score 64; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.3e-54;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 109 EMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHLEQEKLSPL 168
      |||
Db 1 EMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHLEQEKLSPL 60
```

```
QY 169 GEEM 172
      |||
Db 61 GEEM 64
```

```
RESULT 8
US-08-292-870-2
Sequence 2, Application US/08292870
Patent No. 5814467
GENERAL INFORMATION:
APPLICANT: Curtiss, Linda K
APPLICANT: Banka, Carole L
APPLICANT: Bonnet, David J
APPLICANT: Smith, Richard S
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,870
FILING DATE: 17-AUG-1994
```

```
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/534,761
FILING DATE: 07-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/711,333
FILING DATE: 06-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 91/04038
FILING DATE: 07-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-292-870-2
```

```
Query Match      16.5%; Score 44; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.9e-35;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 107 RQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVEPLRAEL 150
      |||
Db 1 RQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKVEPLRAEL 44
```

```
RESULT 9
US-08-292-870-3
Sequence 3, Application US/08292870
Patent No. 5814467
GENERAL INFORMATION:
APPLICANT: Curtiss, Linda K
APPLICANT: Banka, Carole L
APPLICANT: Bonnet, David J
APPLICANT: Smith, Richard S
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,870
FILING DATE: 17-AUG-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/534,761
FILING DATE: 07-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/711,333
FILING DATE: 06-JUN-1991
PRIOR APPLICATION DATA:
```



APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11  
OTHER INFORMATION: /note="Xaa can be either E (Glu)  
or F (Phe)"  
US-08-292-870-3

Query Match 7.9%; Score 21; DB 2; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.4e-13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 QKKQOEEMELYRQKVEPLRAE 149  
Db 12 QKKQOEEMELYRQKVEPLRAE 32

RESULT 10  
US-08-292-870-4  
Sequence 4, Application US/08292870  
Patent No. 5814467  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Linda K  
APPLICANT: Banka, Carole L  
APPLICANT: Bonnet, David J  
APPLICANT: Smith, Richard S  
TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS  
TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/292,870  
FILING DATE: 17-AUG-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/534,761  
FILING DATE: 07-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/711,333  
FILING DATE: 06-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US 91/04038  
FILING DATE: 07-JUN-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 210.1 D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-292-870-4

Query Match 6.4%; Score 17; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SKDLEEVAKAKVQPYLDD 127  
Db 1 SKDLEEVAKAKVQPYLDD 17

RESULT 11  
US-07-959-946-5  
Sequence 5, Application US/07959946  
Patent No. 5408038  
GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Wiltzium, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 No. 5408038th Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/959,946  
FILING DATE: 19921008  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,706  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamsen, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-959-946-5

Query Match 6.0%; Score 16; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KVQPYLDDFQKKWQEE 135  
|||||  
Db 1 KVQPYLDDFQKKWQEE 16

## RESULT 12

US-08-333-577-5  
; Sequence 5, Application US/08333577  
; Patent No. 5786206  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witztum, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5786206th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,577  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER: SCRF 234.0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-333-577-5

Query Match 6.0%; Score 16; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KVQPYLDDFQKKWQEE 135  
|||||  
Db 1 KVQPYLDDFQKKWQEE 16

RESULT 13  
PCT-US92-08634-5  
; Sequence 5, Application PC/TUS9208634  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Witztum, Joseph L.

APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 North Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/08634  
; FILING DATE: 19921009  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US92-08634-5

Query Match 6.0%; Score 16; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 KVQPYLDDFQKKWQEE 135  
|||||  
Db 1 KVQPYLDDFQKKWQEE 16

RESULT 14  
PCT-US94-01234-47  
; Sequence 47, Application PC/TUS9401234  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE  
; TITLE OF INVENTION: BINDING SITES  
; NUMBER OF SEQUENCES: 76  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01234  
; APPLICATION NUMBER: US 08/084,542  
; FILING DATE: 28-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/012,566  
; FILING DATE: 02-FEB-1993  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
PCT-US94-01234-47

Query Match 3.48; Score 9; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 DFQKKWQE 135  
Db 5 DFQKKWQE 13

RESULT 15  
US-08-630-645-6

Sequence 6, Application US/08630645  
Patent No. 5948763

GENERAL INFORMATION:

APPLICANT: SOTO-JARA, Claudio

APPLICANT: BAUMANN, Marc

APPLICANT: FRANGIONE, Blas

TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL COMPOSITIONS

TITLE OF INVENTION: THEREOF FOR TREATMENT OF DISORDERS OR DISEASES ASSOCIATED

TITLE OF INVENTION: WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-LIKE DEPOSITS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 400

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,645

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/478,326

FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C.

REGISTRATION NUMBER: 37,971

REFERENCE/DOCKET NUMBER: SOTO-JARA=1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-630-645-6

Query Match 3.0%; Score 8; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DLATVYVD 44  
Db 1 DLATVYVD 8

RESULT 16  
PCT-US96-10220-6

Sequence 6, Application PC/TUS9610220

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL COMPOSITIONS

TITLE OF INVENTION: THEREOF FOR TREATMENT OF DISORDERS OR DISEASES ASSOCIATED

TITLE OF INVENTION: WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-LIKE DEPOSITS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 400

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/10220

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/478,326

FILING DATE: 06-JUN-1995

APPLICATION NUMBER: US 08/630,645

FILING DATE: 10-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BROWDY, Roger L.

REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: SOTO-JARA=1 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-737-3528

TELEFAX: 202-628-5197

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US96-10220-6

Query Match 3.0%; Score 8; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DLATVYVD 44  
Db 1 DLATVYVD 8

RESULT 17  
US-08-949-155-6

Sequence 6, Application US/08949155

Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A

APPLICANT: Bazer, Fuller W

TITLE OF INVENTION: Compositions and Methods for the

TITLE OF INVENTION: Generation of Transgenic Animal Species

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE AND DURKEE

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210-4433

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/949,155  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/027,338  
FILING DATE: 11-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,094  
FILING DATE: 09-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hibler, David W.  
REGISTRATION NUMBER: 41,071  
REFERENCE/DOCKET NUMBER: TAMK:177  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 317 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-949-155-6

Query Match 3.0%; Score 8; DB 4; Length 317;  
Best Local Similarity 100.0%; Pred. No. 7.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 QLGPTQ 94  
Db 98 QLGPTQ 105

RESULT 18  
US-08-952-796-11  
Sequence 11, Application US/08952796  
Patent No. 6258596  
GENERAL INFORMATION:  
APPLICANT: BENOIT, Patrick  
APPLICANT: BRUCKERT, Eric  
APPLICANT: DENEFFLE, Patrice  
APPLICANT: DUBERGER, Nicolas  
APPLICANT: FRUCHART, Jean-Charles  
APPLICANT: LUC, Gerald  
APPLICANT: TURPIN, Gerard  
APPLICANT: ASSMANN, Gerd  
APPLICANT: FUNKE, Harald  
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPROTEIN A-I  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,796  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 95/06061  
FILING DATE: 22-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00747  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehner Esq., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: ST95031-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-952-796-11

Query Match 2.6%; Score 7; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 MRDRARA 178  
Db 1 MRDRARA 7

RESULT 19  
US-08-448-606-2  
Sequence 2, Application US/08448606  
Patent No. 572114  
GENERAL INFORMATION:  
APPLICANT: Abrahams n, Lars  
APPLICANT: Holmgren, Erik  
APPLICANT: Kalder n, Christina  
APPLICANT: Lake, Mats  
APPLICANT: Mikaelsson, sa  
APPLICANT: Sejlitz, Torsten  
TITLE OF INVENTION: Expression System for Producing  
TITLE OF INVENTION: Apolipoprotein AI-M  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pollock, Vande Sande & Priddy  
STREET: 1990 M Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,606  
FILING DATE: 25-AUG-1995  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE93/01061  
FILING DATE: 09-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9203753-0  
FILING DATE: 11-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Amernick, Burton A.  
REGISTRATION NUMBER: 24,852  
REFERENCE/DOCKET NUMBER: 0151/00121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)331-7111  
TELEFAX: (202) 293-6229

;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 13 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-448-606-2

Query Match 2.6%; Score 7; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DEPPQSP 31  
|||||  
Db 7 DEPPQSP 13

RESULT 20

US-09-101-146-4  
; Sequence 4, Application US/09101146  
; Patent No. 6124125

;; GENERAL INFORMATION:

;; APPLICANT: Dartmouth College, St. Vincents Institute of  
;; APPLICANT: Medical Research, Kemp et al.

;; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase

;; NUMBER OF SEQUENCES: 64

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Jane Massey Licata, Esq.

;; STREET: 66 E. Main Street

;; CITY: Marlton

;; STATE: NJ

;; COUNTRY: USA

;; ZIP: 08053

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

;; COMPUTER: IBM PC

;; OPERATING SYSTEM: WINDOWS 95

;; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/101,146

;; FILING DATE: October 7, 1998

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PN7450

;; FILING DATE: 8 JAN 1996

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Jane Massey Licata

;; REGISTRATION NUMBER: 32,257

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (856) 810-1515

;; TELEFAX: (856) 810-1454

;; INFORMATION FOR SEQ ID NO: 4:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 70

;; TYPE: Amino acid

;; TOPOLOGY: Linear

US-09-101-146-4

Query Match 2.6%; Score 7; DB 3; Length 70;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 VTSTFSK 83  
|||||  
Db 26 VTSTFSK 32

RESULT 21

US-09-188-930-315

; Sequence 315, Application US/09188930A

;; Patent No. 6150502  
;; GENERAL INFORMATION:  
;; APPLICANT: Watson, James D.  
;; APPLICANT: Strachan, Lorna  
;; APPLICANT: Sleeman, Matthew  
;; APPLICANT: Onrust, Rene  
;; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
;; TITLE OF INVENTION: and Methods For Their Use  
;; FILE REFERENCE: 11000.1011c1  
;; CURRENT APPLICATION NUMBER: US/09/188,930A  
;; CURRENT FILING DATE: 1998-11-09  
;; NUMBER OF SEQ ID NOS: 348  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 315  
;; LENGTH: 229  
;; TYPE: PRT  
;; ORGANISM: Rat  
US-09-188-930-315

Query Match 2.6%; Score 7; DB 4; Length 229;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KDLEEVK 118  
|||||  
Db 10 KDLEEVK 16

RESULT 22

US-09-188-930-139

; Sequence 139, Application US/09188930A

; Patent No. 6150502

;; GENERAL INFORMATION:

;; APPLICANT: Watson, James D.

;; APPLICANT: Strachan, Lorna

;; APPLICANT: Sleeman, Matthew

;; APPLICANT: Onrust, Rene

;; TITLE OF INVENTION: Compositions Isolated From Skin Cells

;; TITLE OF INVENTION: and Methods For Their Use

;; FILE REFERENCE: 11000.1011c1

;; CURRENT APPLICATION NUMBER: US/09/188,930A

;; CURRENT FILING DATE: 1998-11-09

;; NUMBER OF SEQ ID NOS: 348

;; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 139

;; LENGTH: 233

;; TYPE: PRT

;; ORGANISM: Rat

US-09-188-930-139

Query Match 2.6%; Score 7; DB 4; Length 233;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 KDLEEVK 118  
|||||  
Db 10 KDLEEVK 16

RESULT 23

US-09-180-827-6

; Sequence 6, Application US/09180827

; Patent No. 6355464

;; GENERAL INFORMATION:

;; APPLICANT: Healy, Judith M.

;; APPLICANT: Bodorova, Jana

;; APPLICANT: Lam, Kelvin T.

;; APPLICANT: Lesson, Andrea J.

;; TITLE OF INVENTION: M. Tuberculosis RNA Polymerase Alpha



TITLE OF INVENTION: Subunit  
FILE REFERENCE: 0342/1C382-US2  
CURRENT APPLICATION NUMBER: US/09/180,827  
CURRENT FILING DATE: 1999-01-26  
PRIOR APPLICATION NUMBER: PCT/US97/22216  
PRIOR FILING DATE: 1997-11-03  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 6  
LENGTH: 315  
TYPE: PRT  
ORGANISM: B. subtilis  
US-09-180-827-6

Query Match 2.6%; Score 7; DB 4; Length 315;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 LEEVKAK 120  
Db 296 LEEVKAK 302

## RESULT 24

US-08-336-891-2  
Sequence 2, Application US/08336891  
Patent No. 5622842  
GENERAL INFORMATION:  
APPLICANT: HOLLIS, GREGORY F.  
APPLICANT: PATEL, MAYUR D.  
TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN A  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHRISTINE E. CARTY  
STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,891  
FILING DATE: 09-NOV-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E.  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-6734  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 343 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-336-891-2

Query Match 2.6%; Score 7; DB 1; Length 343;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 KPALEDL 238  
Db 122 KPALEDL 128

## RESULT 25

PCT-US95-13795-4  
Sequence 4, Application PC/TUS9513795  
GENERAL INFORMATION:  
APPLICANT: HOLLIS, GREGORY F.  
APPLICANT: PATEL, MAYUR D.  
TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHRISTINE E. CARTY  
STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13795  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E.  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19211Y  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-6734  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 343 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-13795-4

Query Match 2.6%; Score 7; DB 5; Length 343;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 KPALEDL 238  
Db 122 KPALEDL 128

## RESULT 26

US-08-637-823B-32  
Sequence 32, Application US/08637823B  
Patent No. 6184031  
GENERAL INFORMATION:  
APPLICANT: Gros, Philippe  
APPLICANT: Skamene, Emil  
TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL  
TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: 411 Hackensack Ave  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,823B  
FILING DATE: 05/08/96  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson, David A.  
REGISTRATION NUMBER: 26,742  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487 5800  
TELEFAX: 201 343 1684  
TELEX:  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 482 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-637-823B-32

Query Match 2.6%; Score 7; DB 4; Length 482;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAARL 202  
Db 40 QRLAARL 46

RESULT 27  
US-08-637-823B-4  
Sequence 4, Application US/08637823B  
Patent No. 6184031  
GENERAL INFORMATION:  
APPLICANT: Gros, Philippe  
TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL  
TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: 411 Hackensack Ave  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,823B  
FILING DATE: 05/08/96  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson, David A.  
REGISTRATION NUMBER: 26,742  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487 5800  
TELEFAX: 201 343 1684  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 483 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-637-823B-4

Query Match 2.6%; Score 7; DB 4; Length 483;

Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 196 QRLAARL 202  
Db 40 QRLAARL 46

RESULT 28  
US-08-637-823B-2  
Sequence 2, Application US/08637823B  
Patent No. 6184031  
GENERAL INFORMATION:  
APPLICANT: Gros, Philippe  
TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL  
TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: 411 Hackensack Ave  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,823B  
FILING DATE: 05/08/96  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson, David A.  
REGISTRATION NUMBER: 26,742  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487 5800  
TELEFAX: 201 343 1684  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-637-823B-2

Query Match 2.6%; Score 7; DB 4; Length 484;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAARL 202  
Db 40 QRLAARL 46

RESULT 29  
US-08-903-139B-7  
Sequence 7, Application US/08903139B  
Patent No. 6114118  
GENERAL INFORMATION:  
APPLICANT: Joe W. Templeton, Jianwei Feng, L. Garry Adams,  
APPLICANT: Erwin Schurr, Philippe Gros, Donald S. Davis and Roger Smith  
TITLE OF INVENTION: METHOD OF IDENTIFICATION OF ANIMALS  
TITLE OF INVENTION: RESISTANT OR SUSCEPTIBLE TO DISEASES SUCH AS RUMINANT  
TITLE OF INVENTION: BRUCELLOSIS, TUBERCULOSIS, PARATUBERCULOSIS AND SALMONELLOSIS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pravel, Hewitt, Kimball & Krieger  
STREET: 1177 West Loop South, 10th Floor

CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77027-9095  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/903,139B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/031,443  
FILING DATE: September 20, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Krieger, Paul E.  
REGISTRATION NUMBER: 25,886  
REFERENCE/DOCKET NUMBER: 00162-3/V96171US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-850-0909  
TELEFAX: 713-850-0165  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 547 amino acids  
TYPE: amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-903-139B-7

Query Match 2.6%; Score 7; DB 3; Length 547;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAARL 202  
Db 104 QRLAARL 110

RESULT 30  
US-08-676-279-50  
Sequence 50, Application US/08676279  
Patent No. 5869247  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: MACROPHAGE NUCLEOTIDE SEQUENCE  
NUMBER OF SEQUENCES: 63  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,279  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/00095  
APPLICATION NUMBER: GB 9400929.7  
FILING DATE: 19-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9422021.7  
FILING DATE: 31-OCT-1994  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 548 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-676-279-50

Query Match 2.6%; Score 7; DB 2; Length 548;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAARL 202  
Db 104 QRLAARL 110

RESULT 31  
US-08-903-139B-8  
Sequence 8, Application US/08903139B  
Patent No. 6114118  
GENERAL INFORMATION:  
APPLICANT: Joe W. Templeton, Jianwei Feng, L. Garry Adams,  
APPLICANT: Erwin Schurr, Philippe Gros, Donald S. Davis and Roger Smith  
TITLE OF INVENTION: METHOD OF IDENTIFICATION OF ANIMALS  
TITLE OF INVENTION: RESISTANT OR SUSCEPTIBLE TO DISEASES SUCH AS RUMINANT  
TITLE OF INVENTION: BRUCELOSIS, TUBERCULOSIS, PARATUBERCULOSIS AND SALMONELLOSIS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESS: Pravel, Kimball & Krieger  
STREET: 1177 West Loop South, 10th floor  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77027-9095  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/903,139B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/031,443  
FILING DATE: September 20, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Krieger, Paul E.  
REGISTRATION NUMBER: 25,886  
REFERENCE/DOCKET NUMBER: 00162-3/V96171US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-850-0909  
TELEFAX: 713-850-0165  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 548 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-903-139B-8

Query Match 2.6%; Score 7; DB 3; Length 548;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 QRLAARL 202  
Db 104 QRLAARL 110

RESULT 32  
US-08-903-139B-9  
Sequence 9, Application US/08903139B  
Patent No. 6114118  
GENERAL INFORMATION:  
APPLICANT: Joe W. Templeton, Jianwei Feng, L. Garry Adams,  
APPLICANT: Erwin Schurr, Philippe Gros, Donald S. Davis and Roger Smith  
TITLE OF INVENTION: METHOD OF IDENTIFICATION OF ANIMALS  
TITLE OF INVENTION: RESISTANT OR SUSCEPTIBLE TO DISEASES SUCH AS RUMINANT

;; TITLE OF INVENTION: BRUCELLSIS, TUBERCULOSIS, PARATUBERCULOSIS AND SALMONELLOSIS  
;; NUMBER OF SEQUENCES: 31  
;; CORRESPONDENCE ADDRESSES:  
;; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger  
;; STREET: 1177 West Loop South, 10th Floor  
;; CITY: Houston  
;; STATE: TX  
;; COUNTRY: USA  
;; ZIP: 77027-9095  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/903,139B  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/031,443  
;; FILING DATE: September 20, 1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Krieger, Paul E.  
;; REGISTRATION NUMBER: 25,886  
;; REFERENCE/DOCKET NUMBER: 00162-3/V96171US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 713-850-0909  
;; TELEFAX: 713-850-0165  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 548 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-903-139B-9

OY 196 QRLAARL 202  
Db 104 QRLAARL 110

Query Match 2.6%; Score 7; DB 3; Length 548;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 33  
US-08-903-139B-28  
; Sequence 28, Application US/08903139B  
; Patent No. 6114118  
; GENERAL INFORMATION:  
; APPLICANT: Joe W. Templeton, Jianwei Feng, L. Garry Adams,  
; TITLE OF INVENTION: METHOD OF IDENTIFICATION OF ANIMALS  
; TITLE OF INVENTION: RESISTANT OR SUSCEPTIBLE TO DISEASES SUCH AS RUMINANT  
; TITLE OF INVENTION: BRUCELLSIS, TUBERCULOSIS, PARATUBERCULOSIS AND SALMONELLOSIS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger  
; STREET: 1177 West Loop South, 10th Floor  
; CITY: Houston  
; STATE: TX  
; COUNTRY: USA  
; ZIP: 77027-9095  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/903,139B  
; FILING DATE:  
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/031,443  
;; FILING DATE: September 20, 1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Krieger, Paul E.  
;; REGISTRATION NUMBER: 25,886  
;; REFERENCE/DOCKET NUMBER: 00162-3/V96171US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 713-850-0909  
;; TELEFAX: 713-850-0165  
;; INFORMATION FOR SEQ ID NO: 28:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 548 amino acids  
;; TYPE: peptide  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-903-139B-28

OY 196 QRLAARL 202  
Db 104 QRLAARL 110

Query Match 2.6%; Score 7; DB 3; Length 548;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 34  
US-08-637-823B-25  
; Sequence 25, Application US/08637823B  
; Patent No. 6184031  
; GENERAL INFORMATION:  
; APPLICANT: Gros, Philippe  
; TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL  
; TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: KLAUBER & JACKSON  
; STREET: 411 Hackensack Ave  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/637,823B  
; FILING DATE: 05/08/96  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson, David A.  
; REGISTRATION NUMBER: 26,742  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487 5800  
; TELEFAX: 201 343 1684  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 548 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-637-823B-25

Query Match 2.6%; Score 7; DB 4; Length 548;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 QRLAARL 202  
1111111

Db 104 QRLAARL 110

## RESULT 35

US-08-676-279-57

Sequence 57, Application US/08676279

Patent No. 5869247

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: MACROPHAGE NUCLEOTIDE SEQUENCE

NUMBER OF SEQUENCES: 63

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/676, 279

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB95/00095

APPLICATION NUMBER: GB 9400929.7

FILING DATE: 19-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9422021.7

FILING DATE: 31-OCT-1994

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 549 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-676-279-57

Query Match 2.6%; Score 7; DB 2; Length 549;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 QRLAARL 202  
1111111

Db 106 QRLAARL 112

## RESULT 36

US-08-637-823B-30

Sequence 30, Application US/08637823B

Patent No. 6184031

GENERAL INFORMATION:

APPLICANT: Gros, Philippe

APPLICANT: Skamene, Emil

TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL

TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: 411 Hackensack Ave

CITY: Hackensack

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/637, 823B

FILING DATE: 05/08/96

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson, David A.

REGISTRATION NUMBER: 26,742

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487 5800

TELEFAX: 201 343 1684

TELEX:

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 550 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-637-823B-30

Query Match 2.6%; Score 7; DB 4; Length 550;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 QRLAARL 202  
1111111

Db 107 QRLAARL 113

## RESULT 37

US-08-637-823B-27

Sequence 27, Application US/08637823B

Patent No. 6184031

GENERAL INFORMATION:

APPLICANT: Gros, Philippe

APPLICANT: Skamene, Emil

TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL

TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: 411 Hackensack Ave

CITY: Hackensack

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/637, 823B

FILING DATE: 05/08/96

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson, David A.

REGISTRATION NUMBER: 26,742

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487 5800

TELEFAX: 201 343 1684

TELEX:

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 568 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-637-823B-27

Query Match 2.6%; Score 7; DB 4; Length 568;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 QRLAARL 202



Db 119 ORLAARL 125

## RESULT 38

US-08-637-823B-28

; Sequence 28, Application US/08637823B  
; Patent No. 6184031

## GENERAL INFORMATION:

APPLICANT: Gros, Philippe

APPLICANT: Skamene, Emil

TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL

TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER &amp; JACKSON

STREET: 411 Hackensack Ave

CITY: Hackensack

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/637,823B

FILING DATE: 05/08/96

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson, David A.

REGISTRATION NUMBER: 26,742

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487 5800

TELEFAX: 201 343 1684

TELEX:

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 584 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Mus musculus

TISSUE TYPE: Blood

CELL TYPE: B-cell precursor

CELL LINE: 70/Z

IMMEDIATE SOURCE:

CLONE: M1ramp-2

US-08-637-823B-28

Query Match

Best Local Similarity 2.6%; Score 7; DB 4; Length 584;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 ORLAARL 202

Db 135 ORLAARL 141

RESULT 39

US-09-300-909-19

; Sequence 19, Application US/09300909  
; Patent No. 6306580

## GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING

TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR

NUMBER OF SEQUENCES: 27

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/300,909

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/083,942

FILING DATE: 01-MAY-1998

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 629 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-300-909-19

Query Match

Best Local Similarity 2.6%; Score 7; DB 4; Length 629;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 VSFLSAL 257

Db 434 VSFLSAL 440

RESULT 40

US-09-413-814-78

; Sequence 78, Application US/09413814  
; Patent No. 6225064

## GENERAL INFORMATION:

APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH

APPLICANT: Bristol-Myers Squibb, Co.

APPLICANT: Beyer, Stefan

APPLICANT: Bloeker, Helmut

APPLICANT: Brandt, Petra

APPLICANT: Cino, Paul M

APPLICANT: Dougherty, Brian A

APPLICANT: Goldberg, Steven L

APPLICANT: Hoffe, Gerhard

APPLICANT: Mueller, Joachim

APPLICANT: Reichenbach, Hans

TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or

TITLE OF INVENTION: heteropolysaccharide compounds

FILE REFERENCE: PCT/US 99/23535

CURRENT APPLICATION NUMBER: US/09/413,814

CURRENT FILING DATE: 1999-10-07

EARLIER APPLICATION NUMBER: DE 198 46 493.2

EARLIER FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 107

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 78

LENGTH: 882

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-413-814-78

Query Match

Best Local Similarity 2.6%; Score 7; DB 4; Length 882;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 LAARLEA 204

Db 158 LAARLEA 164

RESULT 41

US-08-677-010-3

; Sequence 3, Application US/08677010  
; Patent No. 5925805

## GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING

TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR

NUMBER OF SEQUENCES: 27

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
;; GENERAL INFORMATION:
;; APPLICANT: Ohlrogge, John B.
;; APPLICANT: Roesler, Keith R.
;; APPLICANT: Shorosh, Basil S.
;; TITLE OF INVENTION: Structure and Expression of an
;; TITLE OF INVENTION: Arabidopsis Acetyl-coenzyme A Carboxylase Gene
;; NUMBER OF SEQUENCES: 3
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
;; STREET: P.O. Box 828
;; CITY: Bloomfield Hills
;; STATE: Michigan
;; COUNTRY: U.S.A.
;; ZIP: 48303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/677,010
;; FILING DATE:
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Dean F.
;; REGISTRATION NUMBER: 36683
;; REFERENCE/DOCKET NUMBER: 6550-00002CPA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810)641-1600
;; TELEFAX: (810)641-0270
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2254 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; US-08-677-010-3

Query Match          2.6%; Score 7; DB 2; Length 2254;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 RQRLAR 201
   |||||
Db 1501 RQRLAR 1507

RESULT 42
US-08-790-519-3
; Sequence 3, Application US/08790519
; Patent No. 5962767
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Roesler, Keith R.
; APPLICANT: Shorosh, Basil S.
; TITLE OF INVENTION: Structure and Expression of an
; TITLE OF INVENTION: Arabidopsis Acetyl-coenzyme A Carboxylase Gene
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
;; APPLICATION NUMBER: US/08/790,519
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/248,630
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Dean F.
;; REGISTRATION NUMBER: 36683
;; REFERENCE/DOCKET NUMBER: 6550-00002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810)641-1600
;; TELEFAX: (810)641-0270
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2254 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; US-08-790-519-3
```

```
Query Match          2.6%; Score 7; DB 2; Length 2254;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 RQRLAR 201
   |||||
Db 1501 RQRLAR 1507
```

```
RESULT 43
US-08-469-005A-10
; Sequence 10, Application US/08469005A
; Patent No. 5665874
; GENERAL INFORMATION:
; APPLICANT: KUHAJDA, FRANCIS P.
; APPLICANT: PASTERNAK, GARY A.
; TITLE OF INVENTION: CANCER RELATED ANTIGEN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,005A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/188,426
; FILING DATE: 24-JAN-1994
; APPLICATION NUMBER: 08/096,908
; FILING DATE: 26-JUL-1993
; APPLICATION NUMBER: 07/917,716
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 062482-0113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
```

SEQUENCE CHARACTERISTICS:  
LENGTH: 2509 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE: US-08-469-005A-10

Query Match 2.6%; Score 7; DB 1; Length 2509;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 QEFWDNL 99  
|||||  
Db 20 QEFWDNL 26

RESULT 44  
US-09-261-907-2  
Sequence 2, Application US/09261907A  
Patent No. 6294364  
GENERAL INFORMATION:  
APPLICANT: ELLIS, CATHERINE  
APPLICANT: LONSDALE, JOHN  
APPLICANT: BERGSMÄ, DEK J.  
APPLICANT: MOONEY, JEFFREY L.  
APPLICANT: DEPIERA, MEGAN E.  
APPLICANT: CHAPMAN, CONRAD  
TITLE OF INVENTION: HUMAN FAS  
FILE REFERENCE: GP-70603  
CURRENT APPLICATION NUMBER: US/09/261,907A  
CURRENT FILING DATE: 1999-03-03  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 2511  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
US-09-261-907-2

Query Match 2.6%; Score 7; DB 4; Length 2511;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 QEFWDNL 99  
|||||  
Db 20 QEFWDNL 26

RESULT 45  
US-09-144-085-1  
Sequence 1, Application US/09144085  
Patent No. 6280999  
GENERAL INFORMATION:  
APPLICANT: Gustafsson, Claes  
APPLICANT: Betlach, Mary C.  
APPLICANT: Ashley, Gary  
APPLICANT: Julien, Bryan  
APPLICANT: Ziermann, Rainer  
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 30062-20020.20  
CURRENT APPLICATION NUMBER: US/09/144,085  
CURRENT FILING DATE: 1998-08-31  
EARLIER APPLICATION NUMBER: 09/010,809  
EARLIER FILING DATE: 1998-01-22  
NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 5087  
TYPE: PRT  
ORGANISM: Sorangium cellulosum  
US-09-144-085-1

Query Match 2.6%; Score 7; DB 4; Length 5087;  
Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 GGARLAE 215  
|||||  
Db 4467 GGARLAE 4473

RESULT 46  
US-08-704-170-43  
Sequence 43, Application US/08704170  
Patent No. 5707626  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angelina  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 No. 5707626th Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704,170  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-43

Query Match 2.2%; Score 6; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 SKLREQ 87  
|||||  
Db 1 SKLREQ 6

RESULT 47

US-08-474-741-1

; Sequence 1, Application US/08474741  
; Patent No. 5837816  
; GENERAL INFORMATION:  
; APPLICANT: Ciardelli, Thomas L.  
; APPLICANT: Johnson, Kirk  
; TITLE OF INVENTION: METHODS OF PREPARING SOLUBLE, OLIGOMERIC  
; PROTEINS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street - R440  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94608-2916

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,741  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Grant D.  
; REGISTRATION NUMBER: 31,259  
; REFERENCE/DOCKET NUMBER: 1095.002

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2706  
; TELEFAX: (510) 655-3542  
; TELEX: n/a

INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-474-741-1

Query Match 2.2%; Score 6; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 LEALKE 207  
Db 1 LEALKE 6

RESULT 48  
US-09-187-859-1184  
; Sequence 1184, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1184  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-6 cell adhesion recognition sequence

US-09-187-859-1184

Query Match 2.2%; Score 6; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LEEYTK 262  
Db 2 LEEYTK 7

RESULT 49  
PCT-US94-02631-43  
; Sequence 43, Application PC/TUS9402631  
; GENERAL INFORMATION:

; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; TITLE OF INVENTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 North Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/02631  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/029,850  
; FILING DATE: 11-MAR-1993

ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-331  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003

INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US94-02631-43

Query Match 2.2%; Score 6; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 SKLREQ 87  
Db 1 SKLREQ 6

RESULT 50  
US-09-187-859-1147  
; Sequence 1147, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.

;  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1147  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-6 cell adhesion recognition sequence  
US-09-187-859-1147

Query Match 2.28; Score 6; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. NO. 1.7e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 257 LEEYTK 262  
|||||  
Db 3 LEEYTK 8

Search completed: September 22, 2002, 12:23:14  
Job time: 394 sec